(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 September 2003 (25.09.2003)

PCT

(10) International Publication Number
WO 03/078580 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US03/07552

(22) International Filing Date: 13 March 2003 (13.03.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/363,861 13 March 2002 (13.03.2002) US

(71) Applicant *(for all designated States except US)*: PIONEER HI-BRED INTERNATIONAL, INC. [US/US]; 800 Capital Square, 400 Locust Street, Des Moines, IA 50309 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: DANILEVSKAYA, Olga [US/US]; 6004 Dogwood Circle, Johnston, IA 50131 (US). HERMON, Pedro [US/US]; 9814 Newport Vista Drive, Johnston, IA 50131 (US).

(74) Agents: VARLEY, Karen, K. et al.; Darwin Building, 7100 N.W. 62nd Avenue, Johnston, IA 50131-1000 (US).

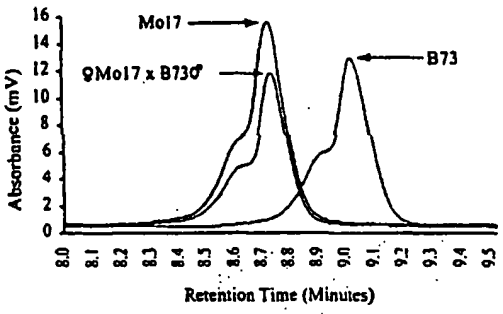
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IMPRINTING IN PLANTS TO CONTROL GENE EXPRESSION



(57) Abstract: Compositions and methods for identifying imprinting and genes regulated by imprinting are provided. The methods involve an analysis of the nucleotide sequence and the identification of CpG islands. At least two islands are involved in imprinting. Thus, genes can be identified that are differentially expressed based on parental inheritance. In this manner, the methods are useful for determining the propensity of a gene to be influenced by imprinting. Such analysis involves determining the pattern of imprinting for cells of interest. It is further recognized that DNA constructs can be constructed which show differential expression depending upon the parent-of-origin. To silence a paternally inherited allele, at least two CpG islands are utilized in the construct.

WO 03/078580 A2

IMPRINTING IN PLANTS TO CONTROL GENE EXPRESSION

BACKGROUND OF THE INVENTION

Genomic imprinting is an epigenetic modification of a specific parental chromosome in the gamete or zygote that leads to monoallelic or differential expression of the two alleles of a gene in somatic cells of the offspring. The general assumption is that maternally- and paternally-transmitted genes are expressed at equivalent levels in progeny. However, non-equivalent expression of the maternally- and paternally-transmitted genes was described in 1970 and 1983 in plants (maize) and mammals, respectively (Alleman M, Plant Mol Biol. (2000) 43:147-61). This phenomenon, named imprinting, is defined as epigenetic gene silencing that is set in the male or female germ lines, resulting in a differential expression of maternally- and paternally-derived alleles. Imprinting affects various essential cellular and developmental processes, including intercellular signaling, RNA processing, cell cycle control, and promotion or inhibition of cellular division and growth.

Many mammalian genes influenced by imprinting have been identified. The first deduction of imprinting at the single gene level involved a transgenic C-myc gene that showed dependence of its expression on paternal inheritance. The silent maternally inherited copy was methylated (Swain *et al.* (1987) *Cell* 50:719-727).

The increased attention to imprinting in mammals is due to the recognition of its importance during development and its role in causing several human genetic diseases. Abnormalities of a single gene can affect imprinting of a proximate genomic region and disrupt multiple disease-causing genes, the phenotype depending upon the parental origin of the mutated gene. Imprinted loci have been implicated in disease. For example, disrupted imprinting of a locus is one of the causes of Prader-Willi syndrome (PWS) and Angelman syndrome (AS), which involve mental retardation. PWS also causes obesity, and AS involves gross motor disturbances. Each disorder can be caused by parental-origin specific uniparental disomy (Nicholls *et al.* (1989) *Nature* 342:281-285; Knoll *et al.* (1990) *Am. J. Hum. Genet.* 47:149-155) or chromosomal deletions (Knoll *et al.* (1989) *Am. J. Hum. Genet.* 47:149-155; Mattei *et al.* (1984) *Hum. Genet.* 66:313-334).

Genomic imprinting has been implicated in cancer. The work has demonstrated that a balance of maternal and paternal chromosomes is required. A relative imbalance leads to neoplastic growth, and the type of neoplasm depends upon whether there is a maternal or paternal genetic excess. Tumors associated with imprinting include the two embryonic tumors, hydatidiform mole and complete ovarian teratoma, familial paraganglioma or glomus tumor, hepatoblastoma

(Rainier *et al.* (1995) *Cancer Res.* 55:1836-1838); (Li *et al.* (1995) *Oncogene* 11:221-229), rhabdomyosarcoma (Zhan *et al.* (1994) *J. Clin. Invest.* 94:445-448), and Ewing's sarcoma (Zhan *et al.* (1995a) *Oncogene* 11:2503-2507). Loss of Imprinting (LOI) of IGF2 and H19 have also now been found in many adult tumors, including uterine (Vu *et al.* (1995) *J. Clin. Endocrinol. Metab.* 80:1670-1676, cervical (Doucrasy *et al.* (1996) *Oncogene* 12:423-430), esophageal (Hibi *et al.* (1996) *Cancer Res.* 56:480-482), prostate (Jarrard *et al.* (1995) *Clin. Cancer Res.* 1:1471-1478), lung cancer (Kondo *et al.* (1995) *Oncogene* 10:1193-1198), choriocarcinoma (Hashimoto *et al.* (1995) *Nat. Genet.* 9:109-110), germ cell tumors (Van Gurp *et al.* (1994) *J. Natl. Cancer Inst.* 86:1070-1075), BWS (Steenman *et al.* (1994) *Nature Genet.* 7:433-439); Weksberg *et al.* (1993) *Nature Genet.* 5:143-150), and Wilms tumor (Ogawa *et al.* (1993) *Nature Genet* 5:408-412). In the case of familial paraganglioma, the transmitting parent is the father (Van der Mey *et al.* (1989) *Lancet* 2:1291-1294). The gene has recently been localized to 11q22.3-q23 (Heutink *et al.* (1994) *Eur. J. Hum. Genet* 2:148-158).

In angiosperm plants, imprinting is postulated to be essential for endosperm development. In *Arabidopsis*, the *MEA* gene regulates cell proliferation by exerting a gametophytic maternal control during seed development. Seeds derived from embryo sacs carrying a mutant *mea*-1 allele abort after delayed morphogenesis with excessive cell proliferation in the embryo and reduced free nuclear divisions in the endosperm. The mutant *mea* seeds are able, at a low frequency, to initiate endosperm development, seed coat differentiation, and fruit maturation in the absence of fertilization. See, Vielle-Calzada *et al.* (1999) *Genes & Development* 13:2971-2982. The *mea* mutation affects an imprinted gene expressed maternally in cells of the female gametophyte and after fertilization only from maternally inherited *MEA* alleles. Paternally inherited *MEA* alleles are transcriptionally silent in both the young embryo and endosperm.

A consequence of imprinting is the requirement of a 2:1 ratio of maternal to paternal genomes in the endosperm (Haig and Westoby 1991, Am. Nat. 134:147-155). Thus imprinting plays a significant role in the proper development of seed in cereal crops.

Abnormal imprinting has been studied in plants by analysis of gene expression. Methods are needed in the art to identify imprinted genes in plants, to identify genes involved in endosperm development, and to manipulate gene sequences to affect imprinting.

BRIEF SUMMARY OF THE INVENTION

Compositions and methods for identifying imprinting and genes regulated by imprinting are provided. The methods involve an analysis of the nucleotide sequence and the identification of CpG islands. At least two islands are involved in imprinting. Thus, genes can be identified that are differentially expressed based on parental inheritance. In this manner, the methods are useful for determining the propensity of a gene to be influenced by imprinting. Such analysis involves determining the pattern of imprinting for cells of interest.

It is further recognized that DNA constructs can be created which show differential expression depending upon the parent of origin. To silence a paternally inherited allele, at least two CpG islands are utilized in the construct.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the massively parallel signature sequencing (MPSS) analysis of *ZmFIE1* expression in embryo and endosperm. The graph represents a distribution of the 17-mer tags (GATCTAGTGTGTGGCTG) in the endosperm and embryo mRNAs generated by MPSS. The recognition site of the restriction enzyme DpnII, used to generate tags, is GATC. A tag sequence is derived from the *ZmFie1* EST (Accession No. AY061964) positioned 112 nt upstream from the polyA tail. The vertical axis represents the frequency of the tags as particles per million (PPM) molecules sequenced on the microbeads. The horizontal axis represents stages of kernel development starting with unfertilized ovules (point "0"), and 8, 12, 21, 25, and 35 days after pollination (DAP). Endosperm and embryos were dissected from kernels. Note that embryo tissues were not dissected from 8 DAP kernels. Squares indicate endosperm; triangles indicate embryos.

Figure 2 shows the pattern of paternal and maternal *ZmFie1* allele expression in developing kernels. The graphs represent a size-dependent separation of the RT-PCR DNA fragments by the WAVE HPLC System. The larger fragments have a longer retention time on the DNASEP cartridge, which results in an accurate quantitative separation of the complex fragment mixture. Total RNA was isolated from 15 DAP kernels of selfed Mo17 and B73 lines and their reciprocal crosses. RT-PCR was performed with primers positioned around 12 nt deletions at 3' UTR in Mo17 background (2A, 2B). The anonymous EST was used as a control for the expression of both maternal and paternal allele in the same samples of RNA (2C).

Figure 3 shows that *ZmFie2* Mo17 and B73 alleles are polymorphic by the MITE insertion at 3' UTR. The position of a common forward primer F (exon 11) and the genotype-specific reverse primers (3' UTR) are shown by arrows. DNA

sequence of the MITE insertion into 3' UTR of the *ZmFie2* B73 allele is shown in 3B. The target site duplication is boxed. The 14 nt terminal inverted repeats are marked by arrowheads.

Figure 4 shows the genomic structure of the *ZmFie* loci. 12 kb genomic segments of the *ZmFie1* (A) and *ZmFie2* (B) regions are shown. The predicted start and stop codons of *ZmFie* coding regions are indicated by ATG and TGA. The positions of nucleotides are relative to the translation start codon ATG. Exons are shown as tall vertical boxes, untranslated regions as shorter boxes, and introns as connecting double lines. The putative transcription and translation start sites are shown as bent arrows. Regions with homology to retrotransposons are stippled. The direct repeats positioned upstream of *ZmFie2* are marked by large arrows.

Figure 5 shows the 5' upstream and coding sequence for the *ZmFie1* gene sequence.

Figure 6 shows the 5' upstream and coding sequence for the *ZmFie2* gene sequence.

Figure 7 shows the distribution of the CpG and CpNpG methylation sites along the *ZmFie* genomic sequences. The graphs present the number of CpG or CpNpG sites per 100 nt. The start and stop codons are indicated by ATG and TGA. The CpG islands are marked by filled rectangles.

Figure 8 shows a phylogenetic tree of plant FIE proteins.

Figure 9 shows the distribution of HpaII restriction sites across the ZmFIE1 and ZmFIE2 genomic sequences.

Figure 10 is a table of primers designed around clusters of HpaII sites to monitor cytosine methylation.

Figure 11 shows single nucleotide polymorphisms (SNPs) present in exon 1 of B73 and Mo17 inbred lines.

DETAILED DESCRIPTION OF THE INVENTION

Imprinting has been observed in eukaryotic cells of plants and mammals (Yoder and Bestor (1996) *Biol. Chem.* 377(10): 605-610). In humans and other mammals, normal imprinting underlies several fundamental cellular and developmental processes; thus, abnormal imprinting patterns are implicated in a wide variety of catastrophic human diseases. "Imprinting" is defined as an epigenetic modification of a specific parental allele of a gene, or the chromosome on which it resides, in the gamete or zygote, leading to differential expression of the two alleles in somatic cells of the offspring. That is, genomic imprinting is an epigenetic chromosomal modification in the germ line that leads to preferential expression of one of the two parental alleles in a parent-of-origin-specific manner.

"Normal pattern of imprinting" means preferential expression of a single parental allele of an imprinted gene and/or preferential methylation of a single parental allele of an imprinted gene. "Loss of imprinting" or "LOI" means loss of a normal pattern of imprinting, i.e., the loss of preferential expression of a single parental allele of an imprinted gene and/or the loss of methylation of a single parental allele of an imprinted gene. LOI is exhibited by a variety of abnormal expression patterns. Such patterns include but are not limited to: equal expression of both alleles; significant (>5%) expression of the normally silent allele when the normal case is complete silencing of one allele; epigenetic silencing of the normally expressed copy of an imprinted gene; the absence of methylation of both alleles and/or the methylation of both alleles where the normal case is methylation of a single allele.

Imprinting is a developmental phenomenon wherein a gene in a gamete or zygote is modified such that preferential expression of a single parental allele occurs in the offspring. It has been theorized that "CpG islands" present within the gene are subject to methylation, which causes repression of one allele (Stoger *et al.* (1993) *Cell* 73:61-71). CpG islands are defined as sequences of 200 or more base pairs with a GC content greater than 0.5 and an observed-to-expected CpG dinucleotide content greater than 0.6 (Gardiner-Garden and Frommer (1987) *J. Mol. Biol.* 196:261-282). Allele-specific methylation of CpG islands is a feature of the inactive X chromosome (Yen *et al.* (1984) *Proc. Natl. Acad. Sci. USA* 81:1759-1763) and imprinted genes including *H19*, *Snrpn*, and *Igf2r* (Brandeis *et al.* (1993) *EMBOJ* 12:3669-3677; Shemer *et al.* (1997) *Proc. Natl. Acad. Sci. USA* 94:10267-10272; Wutz *et al.* (1997) *Nature* 389:745-749). Analysis of orthologous genomic domains of approximately 1 Mb in mouse and human identified nine conserved imprinted genes; in eight of these, two or more conserved CpG islands were found upstream of or within the gene. In contrast, six non-imprinted genes within the same region were associated with at most one CpG island (Onyango *et al.* (2000) *Genome Research* 10:1697-1710).

The present invention has identified CpG islands in plants and attributes differential expression of imprinted plant genes to CpG islands. Accordingly, the methods of the invention encompass the identification of imprinted plant genes by determining the presence of CpG islands. Where sequence information is available for a plant, the sequence can be searched for GC rich regions and further testing can be done to establish the location of CpG islands.

Methods for the determination of the pattern of imprinting are known in the art. It is recognized that the methods may vary depending on the gene to be analyzed. Generally, in methods for assaying allele-specific gene expression, RNA is reverse transcribed with reverse transcriptase, and then PCR is performed

with PCR primers that span a site within an exon where that site is polymorphic (i.e., normally variable in the population), and this analysis is performed on an individual that is heterozygous (i.e., informative) for the polymorphism. One then uses any of a number of detection schemes to determine whether one or both alleles is expressed. Methods for the assessment of gene expression, allele-specific gene expression, and DNA methylation are encompassed. Additionally, direct approaches to identifying novel imprinted genes include: positional cloning efforts aimed at identifying imprinted genes near other known imprinted genes (Barlow *et al.* (1991) *Nature* 349:84-87); techniques comparing gene expression (Kuroiwa *et al.* (1996) *Nat. Genet.* 12:186-190); and restriction landmark genome scanning (Nagai *et al.* (1995) *Biochem. Biophys. Res. Commun.* 213:258-265). See also, Rainier *et al.* (1993) *Nature* 362:747-749; which teaches the assessment of allele-specific expression of IGF2 and H19 by reverse-transcribing RNA and amplifying cDNA by PCR using new primers that permit a single round rather than nested PCR; Matsuoka *et al.* (1996) *Proc. Natl. Acad. Sci USA* 93:3026-3030, which teaches the identification of a transcribed polymorphism in $p57^{KIP2}$; Thompson *et al.* (1996) *Cancer Research* 56:5723-5727, which teaches determination of mRNA levels by RPA and RT-PCR analysis of allele-specific expression of $p57^{KIP2}$; and Lee *et al.* (1997) *Nature Genetics* 15:181-185, which teaches RT-PCR SSCP analysis of two polymorphic sites. Such disclosures are herein incorporated by reference.

Direct approaches developed to identify novel imprinted genes include: positional cloning, which identifies imprinted genes near other known imprinted genes (Barlow *et al.* (1991) *Nature* 349:84-87); comparing gene expression in parthenogenetic embryos to that of normal embryos (Kuroiwa *et al.* (1996) *Nat. Genet* 12:186-190); and restriction landmark genome scanning (Nagai *et al.* (1995) *Biochem. Bionhys. Res. Commun.* 213:258-265). The last approach comprises analysis of clonality in tumors by assessing DNA methylation near a heterozygous polymorphic site (Vogelstein *et al.* (1985) *Science* 227:642-645).

As noted above, a distribution of CpG islands within genes can be used as a predictive tool for genes regulated by imprinting. To date, imprinted genes in plants are important components of regulation of endosperm size and growth. Thus, the methods of the invention can be used to identify genes involved in endosperm development. In particular, the invention can be used as a predictive tool for plant genes, dicot and monocot genes, particularly maize genes, that are regulated by imprinting.

It is also recognized that the CpG islands of the invention may be used to silence paternally transmitted genes. In this manner, DNA constructs comprising

at least two CpG islands will be operably linked with a coding sequence and a promoter that is expressed in plants.

A number of promoters can be used in the practice of the invention. The promoters can be selected based on the desired outcome. The nucleic acids can be combined with constitutive, tissue-preferred, or other promoters for expression in plants.

Such constitutive promoters include, for example, the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO 99/43838 and U.S. Patent No. 6,072,050; the core CaMV 35S promoter (Odell *et al.* (1985) *Nature* 313:810-812); rice actin (McElroy *et al.* (1990) *Plant Cell* 2:163-171); ubiquitin (Christensen *et al.* (1989) *Plant Mol. Biol.* 12:619-632 and Christensen *et al.* (1992) *Plant Mol. Biol.* 18:675-689); pEMU (Last *et al.* (1991) *Theor. Appl. Genet.* 81:581-588); MAS (Velten *et al.* (1984) *EMBO J.* 3:2723-2730); ALS promoter (U.S. Patent No. 5,659,026), and the like. Other constitutive promoters include, for example, U.S. Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611.

Chemically-regulated promoters can be used to modulate the expression of a gene in a plant through the application of an exogenous chemical regulator. Depending upon the objective, the promoter may be a chemical-inducible promoter, where application of the chemical induces gene expression, or a chemical-repressible promoter, where application of the chemical represses gene expression. Chemical-inducible promoters are known in the art and include, but are not limited to, the maize In2-2 promoter, which is activated by benzenesulfonamide herbicide safeners, the maize GST promoter, which is activated by hydrophobic electrophilic compounds that are used as pre-emergent herbicides, and the tobacco PR-1a promoter, which is activated by salicylic acid. Other chemical-regulated promoters of interest include steroid-responsive promoters (see, for example, the glucocorticoid-inducible promoter in Schena *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:10421-10425 and McNellis *et al.* (1998) *Plant J.* 14(2):247-257) and tetracycline-inducible and tetracycline-repressible promoters (see, for example, Gatz *et al.* (1991) *Mol. Gen. Genet.* 227:229-237, and U.S. Patent Nos. 5,814,618 and 5,789,156), herein incorporated by reference.

Tissue-preferred promoters can be utilized to target enhanced expression within a particular plant tissue. Tissue-preferred promoters include Yamamoto *et al.* (1997) *Plant J.* 12(2):255-265; Kawamata *et al.* (1997) *Plant Cell Physiol.* 38(7):792-803; Hansen *et al.* (1997) *Mol. Gen Genet.* 254(3):337-343; Russell *et al.* (1997) *Transgenic Res.* 6(2):157-168; Rinehart *et al.* (1996) *Plant Physiol.* 112(3):1331-1341; Van Camp *et al.* (1996) *Plant Physiol.* 112(2):525-535; Canevascini *et al.* (1996) *Plant Physiol.* 112(2):513-524; Yamamoto *et al.* (1994)

Plant Cell Physiol. 35(5):773-778; Lam (1994) *Results Probl. Cell Differ.* 20:181-196; Orozco *et al.* (1993) *Plant Mol Biol.* 23(6):1129-1138; Matsuoka *et al.* (1993) *Proc Natl. Acad. Sci. USA* 90(20):9586-9590; and Guevara-Garcia *et al.* (1993) *Plant J.* 4(3):495-505. Such promoters can be modified, if necessary, for weak expression.

"Seed-preferred" promoters include both "seed-specific" promoters (those promoters active during seed development such as promoters of seed storage proteins) as well as "seed-germinating" promoters (those promoters active during seed germination). See Thompson *et al.* (1989) *BioEssays* 10:108, herein incorporated by reference.

Examples include, for dicotyledonous plants, a bean β-phaseolin promoter, a napin promoter, a β-conglycinin promoter, a cruciferin promoter, and a soybean lectin promoter. For monocotyledonous plants, promoters useful in the practice of the invention include, but are not limited to, cZ19B1 (maize 19 kDa zein), milps (myo-inositol-1-phosphate synthase), celA (cellulose synthase) (see WO 00/11177, herein incorporated by reference), a maize 15 kD zein promoter, a 22 kD zein promoter, a 27Kd γ-zein promoter (such as gzw64A promoter, *see* Genbank Accession #S78780), a waxy promoter, a shrunken-1 promoter, a globulin 1 promoter (See Genbank Accession # L22344), an ltp2 promoter (Kalla, et al., Plant Journal 6:849-860 (1994); U.S. Patent 5,525,716), cim1 promoter (U.S. Patent 6,225,529), maize end1 and end2 promoters (*See* U.S. patent applications 09/383,543, filed August 26, 1999, and 10/310,191, filed December 4, 2002), and the shrunken-2 promoter. See also U.S. patents 6,407,315 and 6,403,862. However, other promoters useful in the practice of the invention are known to those of skill in the art such as nucellain promoter (*See* C. Linnestad, *et al.*, Plant Physiol. 118:1169-80 (1998)), kn1 promoter (*See* S. Hake and N. Ori, B8: INTERACTIONS AND INTERSECTIONS IN PLANT PATHWAYS, COEUR D'ALENE, IDAHO, KEYSTONE SYMPOSIA, February 8-14, 1999, at 27.), and F3.7 promoter (Baszczynski et al., Maydica 42:189-201 (1997)). Spatially acting promoters such as glb1, an embryo-preferred promoter; or gamma zein, an endosperm-preferred promoter, or BETL1 (*See* G. Hueros, *et al.*, Plant Physiology 121:1143-1152 (1999)), are particularly useful. The use of temporally acting promoters is also contemplated by this invention. Promoters that act from 0-25 days after pollination (DAP) are preferred, as are those acting from 4-21, 4-12, or 8-12 DAP. In this regard, promoters such as cim1 and ltp2 are preferred. Particularly preferred promoters include maize zag2.1 (GenBank Accession X80206), maize zap (see U.S. Provisional Patent Application 60/364,065), maize ckx1-2 promoter (see U.S. Patent Publication 2002-0152500 A1), maize end2 (see U.S. Patent

6,528,704, and also U.S. Patent Application 10/310,191, filed December 4, 2002), and maize lec1 (see U.S. Patent Application 09/718,754, filed December 27, 2002).

Transformation protocols as well as protocols for introducing nucleotide sequences into plants may vary depending on the type of plant or plant cell, i.e., monocot or dicot, targeted for transformation. Suitable methods of introducing nucleotide sequences into plant cells and subsequent insertion into the plant genome include microinjection (Crossway *et al.* (1986) *Biotechniques* 4:320-334), electroporation (Riggs *et al.* (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606, *Agrobacterium*-mediated transformation (Townsend *et al.*, U.S. Patent No. 5,563,055; Zhao *et al.*, U.S. Patent No. 5,981,840), direct gene transfer (Paszkowski *et al.* (1984) *EMBO J.* 3:2717-2722), and ballistic particle acceleration (see, for example, Sanford *et al.*, U.S. Patent No. 4,945,050; Tomes *et al.*, U.S. Patent No. 5,879,918; Tomes *et al.*, U.S. Patent No. 5,886,244; Bidney *et al.*, U.S. Patent No. 5,932,782; Tomes *et al.* (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag, Berlin); McCabe *et al.* (1988) *Biotechnology* 6:923-926); and Lec1 transformation (WO 00/28058). Also see Weissinger *et al.* (1988) *Ann. Rev. Genet.* 22:421-477; Sanford *et al.* (1987) *Particulate Science and Technology* 5:27-37 (onion); Christou *et al.* (1988) *Plant Physiol.* 87:671-674 (soybean); McCabe *et al.* (1988) *Bio/Technology* 6:923-926 (soybean); Finer and McMullen (1991) *In Vitro Cell Dev. Biol.* 27P:175-182 (soybean); Singh *et al.* (1998) *Theor. Appl. Genet.* 96:319-324 (soybean); Datta *et al.* (1990) *Biotechnology* 8:736-740 (rice); Klein *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:4305-4309 (maize); Klein *et al.* (1988) *Biotechnology* 6:559-563 (maize); Tomes, U.S. Patent No. 5,240,855; Buising *et al.*, U.S. Patent Nos. 5,322,783 and 5,324,646; Tomes *et al. (1995)* "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg (Springer-Verlag, Berlin) (maize); Klein *et al.* (1988) *Plant Physiol.* 91:440-444 (maize); Fromm *et al.* (1990) *Biotechnology* 8:833-839 (maize); Hooykaas-Van Slogteren *et al.* (1984) *Nature (London)* 311:763-764; Bowen *et al.*, U.S. Patent No. 5,736,369 (cereals); Bytebier *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:5345-5349 (Liliaceae); De Wet *et al.* (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman *et al.* (Longman, New York), pp. 197-209 (pollen); Kaeppler *et al.* (1990) *Plant Cell Reports* 9:415-418 and Kaeppler *et al.* (1992) *Theor. Appl. Genet.* 84:560-566 (whisker-mediated transformation); D'Halluin *et al.* (1992) *Plant Cell* 4:1495-1505 (electroporation); Li *et al.* (1993) *Plant Cell Reports* 12:250-255 and Christou and Ford (1995) *Annals of Botany* 75:407-413 (rice); Osjoda *et al.* (1996)

Nature Biotechnology 14:745-750 (maize via *Agrobacterium tumefaciens*); all of which are herein incorporated by reference.

The cells that have been transformed may be grown into plants in accordance with conventional ways. See, for example, McCormick *et al.* (1986) *Plant Cell Reports* 5:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, and resulting plants having desired expression of the subject phenotypic characteristic may be identified. Two or more generations may be grown to ensure that the desired expression of the subject phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure that desired expression of the subject phenotypic characteristic has been achieved.

The following examples are offered by way of illustration, not by way of limitation.

EXPERIMENTAL

Introduction

A fundamental problem in biology is to understand how fertilization initiates reproductive development. In flowering plants, the female gametophyte, or embryo sac, is composed of egg, central, synergid, and antipodal cells. Double fertilization triggers development of the egg into a diploid embryo and development of the central cell into a triploid endosperm. In sexually-reproducing plants, the embryo sac never develops into seed without fertilization. In asexually-reproducing apomictic plants, the egg cell develops parthenogenetically without fertilization to produce the embryo, but in many species the endosperm development may still require fertilization (non-autonomous apomicts) (Grimanelli *et al.* (2001) *Trends Genet.* 17(10):597-604).

A number of mutants that initiate fertilization independent seed (FIS) development have been isolated in *Arabidopsis* (Ohad *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93(11):5319-5324; Chaudhury *et al.* (1997) *Annu. Rev. Cell Dev. Biol.* 17:677-699). These mutants uncouple seed development from the fertilization process and display some characteristics of apomixis, such as autonomous endosperm development. A mutational approach has revealed three genes with similar FIS phenotypes: FIS1/MEDEA, which is related to the Polycomb group (PcG) protein EZ (enhancer of Zest) of *Drosophila* (Grossniklaus *et al.* (1998) *Science* 280:466-450; Luo *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 94(8):4223-4228); FIS2, which is a C_2H_2 Zinc Finger transcriptional regulator that may have a similar function to Hunch back protein of the *Drosophila* PcG complex (Luo *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 94(8):4223-4228); and FIS3/FIE, which is a homologue of the PcG protein ESC (extra sex combs) (Ohad *et al.*

(1999) *Plant Cell* 11:407-416). Polycomb group proteins are conserved among eukaryotes and are involved in the repression of homeotic genes during early development in flies and mammals. One could speculate that *FIS* genes define a PcG- like complex in plants that suppresses the development of the endosperm in the absence of fertilization (Grossniklaus *et al.* (1998) *Science* 280(5362):466-450; Luo *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 94(8):4223-4228; Ohad *et al.* (1999) *Plant Cell* 11:407-416).

The Arabidopsis model provides candidate genes for revealing similar pathways in other plants. A search of the homologues in a proprietary maize EST (Expressed Sequencing Tags) database identified two maize genes, *ZmFie1* and *ZmFie2* (see WO 01/16325, herein incorporated by reference). The putative FIE maize proteins share 57-68% identity with the *Arabidopsis* FIE protein. *FIS2/3* genes do not demonstrate such a remarkable conservation. A duplication of the maize *Fie* gene raises a question about their functional redundancy. The two *ZmFie* genes show a different pattern of expression in vegetative and reproductive tissues, but they may have overlapping function in the developing kernels.

In this example, the expression of two maize *Fie* genes in developing kernels has been analyzed by several different methods, which lead to the conclusion that the two *ZmFie* genes may have nonredundant functions. Based on the expression pattern and a temporal type of imprinting, *ZmFie2* is likely to be a functional homologue of the *Arabidopsis FIE* gene and most likely is involved in the repression of endosperm development before pollination. The expression of *ZmFie1* is triggered in endosperm after pollination, which implies no repressive function in the embryo sac before pollination, but reveals a new endosperm-specific FIE function in maize. Only the maternal *ZmFIE1* allele is expressed during kernel development, implying a strong regulation by imprinting. Based on the genomic sequences of *ZmFIE* genes, different models for temporal and permanent types of imprinting are proposed. Thus far the *ZmFIE1* gene is found only in maize, which is likely to be a consequence of its allotetraploid origin.

Experimental Procedures

RNA Gel Blot Analysis.

To analyze ZmFIE expression in developing kernels, mRNA was isolated from non-pollinated ovules at silking and from kernels at 3, 6, 9, 12, and 15 days after pollination (DAP). Total RNA was extracted from 1 g of material using a hot phenol extraction procedure and a selective precipitation with 4 M LiCl to remove traces of DNA and small RNA species (Verwoerd *et al.* (1989) *Nucleic Acids Res.* 17:2362; Brugière *et al.* (1999) *Plant Cell* 11:1995-2012). For each time point,

kernels were collected from two ears harvested from two different plants (replications) from either the B73 or Mo17 inbred lines. RNA was quantified using a spectrophotometer at 260 nm. Poly(A) was prepared from total RNA (400 µg) using the Oligotex™ poly(A) purification kit (Qiagen). For gel blot experiments, poly(A) RNA enriched samples were prepared as described by Becker *et al.* (1993) *Methods Enzymol.* 218:568-587. Three µg of polyA RNA were loaded in each lane. Electrophoretic separation was performed on 1.5% agarose gels containing 5% (v/v) of a solution of 37% formaldehyde in Mops buffer (0.02 M Mops, pH 7.0, 5 mM sodium acetate, and 1 mM EDTA). Gels were blotted onto a nylon membrane (Roche Molecular Biochemicals) using TurboBlotter (Schleicher & Schuell), with 20xSSC (1xSSC is 150 mM NaCl, 15 mM sodium citrate) as transfer buffer. Blots were probed with ^{32}P-labeled 300bp fragments of *ZmFIE1* or *ZmFIE2* cut from the 3' UTR of the appropriate ETS clones. The fragment sequences shared no homology, which avoided cross-hybridizations. Actin probe was used as a loading control.

Distinguishing ZmFie mRNAs in Reciprocal Crosses.

Reciprocal crosses between B73 and Mo17 inbred lines were performed, and F1 kernels were sampled at 2, 5, 10, and 15 days after pollination (DAP). Total RNA was isolated and reverse PCR reactions were performed with "Superscript kit." The PCR product differed between B73 and Mo17 alleles in a 12 nt deletion. PCR product was separated on HPLC WAVE machine to distinguish between B73 and Mo17 alleles.

Primers to amplify *ZmFIE2* were designed based on the MITE insertion in the B73 *ZmFIE2* allele. In B73 background, *ZmFIE2* polyA transcripts are terminated in the middle of this insertion. In Mo17 background, *ZmFIE2* polyA transcripts are terminated within genomic sequence with no homology to MITE insertion. (See Figure 3A.) The forward primer positioned in exon eleven, 5'-CGTGAAGGCAAAATCTACGTGTGG-3', (SEQ ID NO: 2) is common for both genotypes. The reverse primer 5'- CATTACGTTACAAATATGTGAACCAAACG-3' (SEQ ID NO: 3) is specific for the B73 allele; reverse primer 5'-CAGAACAAACAGATGACAACGGTTCCCAAAG-3' (SEQ ID NO: 4) is specific for the Mo17 allele. This primer combination allows for monitoring of B73 and Mo17 *ZmFIE2* allele expression in developing kernels of the reciprocal crosses by RT-PCR.

In Situ Hybridization.

To determine expression patterns of *ZmFIE* genes in maize, *in situ* hybridization was performed using the protocol of Jackson (1991) in *In situ*

Hybridization in Plants, Molecular Plant Pathology: A Practical Approach, ed. Bowles *et al.* (Oxford University Press, England), pp. 63-74. Sense and antisense mRNA probes of 300 bp corresponding to the 3' UTR of *ZmFIE* genes were labeled non-isotopically with digoxigenin-UTP by *in vitro* transcription with T7 and T3 RNA polymerases (Roche Molecular Biochemicals). Probes were hybridized with fixed sections of maize tissues from ovules at silking, and kernels at 5, 8, and 12 DAP. Following extensive washing to remove unbound probe, signal was detected with anti-DIG-antibodies conjugated with alkaline phosphatase to mediate color reaction (Roche Molecular Biochemicals) that leads to a purple-blue precipitate in the cells that contain mRNA. *ZmFIE* mRNAs were detected specifically with the antisense probe; the sense probe did not hybridize, therefore serving as a negative control.

Cloning and Sequencing of ZmFIE Genomic Fragments.

BAC genomic libraries were screened with *ZmFIE1* and *ZmFIE2* ESTs. Five BAC clones per each gene were identified and confirmed by Southern hybridization. HindIII and EcoR1 BAC fragments subcloned into vector BluescriptII (KS) (Stratagene) were hybridized with *ZmFIE* probes, and positive clones were sequenced.

DNA Sequence Analysis.

DNA assembly was performed using the Sequencher program (Genecode, Ann Arbor, MI). BLAST search of GenBank was used for sequence annotation. Sequence analysis was performed with GCG® programs (Accelrys, Inc., San Diego, CA).

Nucleotide Sequence Accession Numbers.

The sequences have been deposited in the GenBank database under Accession No. AY061964 (*ZmFie1* genomic locus), and AY061965 (*ZmFie2* genomic locus).

Example 1: Maize *FIE* (Fertilization Independent Endosperm) Homologues: Two Related Genes with Distinct Expression Patterns.

Results

Expression of ZmFIE Genes in Developing Kernels.

ZmFie genes have a different pattern of expression in vegetative and reproductive tissues. Expression of *ZmFIE1* was detected only in developing kernels, not in vegetative tissues. Conversely, *ZmFIE2* expression was found in all tissues tested. If these genes participate in repression of embryo sac development

before fertilization in a manner similar to the *Arabidopsis FIE* homologue, they should be expressed in the ovules before fertilization. To understand the function of both genes, their expression in ovules and developing kernels was detected by mRNA gel blot experiments, gene expression analysis by massively parallel signature sequencing (MPSS) (Brenner et al. (2000) *Nat. Biotechnol.* 18:630-634), and by *in situ* hybridization.

For RNA gel blot experiments, mRNA was isolated from non-pollinated ovules and from developing kernels at 3, 6, 9, 12, and 15 days after pollination (DAP). *ZmFIE1* mRNA is not detected in ovules and 3 DAP kernels. It appears first in 6 DAP kernels, reaching a maximum of expression in 9 DAP kernels, and gradually declines at later stages. The expression pattern of *ZmFIE2* is very different: mRNA is detected in ovules and all stages of developing kernels, but declines after 6 DAP. RNA gel blot experiments demonstrate a low-abundance of *ZmFIE2* mRNA, compared to *ZmFIE1*mRNA, which shows significantly higher expression.

To achieve a more sensitive assay of *ZmFIE* expression, these cDNA sequences were searched with a BLAST algorithm against the gene expression database generated by the MPSS method from different maize tissues. Massively parallel signature sequencing (MPSS) generates 17-mer sequencing tags of millions of cDNA molecules, which are *in vitro* cloned on microbeads (Brenner *et al.* (2000) *Nat. Biotechnol.* 18:630-634). The technique provides an unprecedented depth and sensitivity even for messages that are expressed at very low levels. MPSS is based on the *DpnII* (GATC) restriction site availability in cDNA templates. If the site is absent, the 17-mer tags are not generated. *ZmFie2* does not have the appropriate *DpnII* site and is not suitable for MPSS analysis. For this reason, only *ZmFIE1* tags were found. Distributions of the *ZmFIE1* tags in MPSS experiments are shown in Figure 1. No tags were detected in mRNA isolated from ovules. Thus, if *ZmFIE1* were transcribed in ovules, it would produce less than one mRNA molecule per 10^6 total mRNA molecules. At 8 DAP, the number of tags is about 600 PPM (particles per million), gradually decreasing at later stages and reaching 20 PPM at 35 DAP. No tags are found in 40 DAP kernels. This trend is in complete agreement with mRNA gel blot experiments and RT-PCR (data not shown). The second important observation from MPPS experiments is the expression of *ZmFie1* in the developing endosperm. Embryo and endosperm were dissected for MPSS experiments from kernels as early as 10 DAP. At this stage, *ZmFIE1* expression is approximately 20-30 times higher in endosperm than in embryo. MPSS analysis strongly suggests that transcription of *ZmFIE1* is activated in developing kernels approximately 5-6 days after pollination, predominantly in endosperm.

Because this type of analysis is not available for *ZmFIE2*, *in situ* hybridization was performed. Longitudinal sections of B73 ovules and kernels at 2, 5, 8 and 15 DAP were prepared and hybridized with antisense RNA probes, and with sense RNA probes as a negative control. The sense probe revealed no background signals, and images are not shown. *ZmFIE2* antisense probes gave a signal in the embryo sac of the mature ovules at silking. At 2 DAP, zygotes had a significantly increased signal compared to ovules, indicating that *ZmFIE2* transcription is activated *de novo*, and the signal intensity may not be explained by the pre-existing maternal RNA. In kernels at 5 DAP, the most intense signal appeared in the embryo-surrounding region and on the periphery of the developing endosperm. At the later stage of 15 DAP, the signal persists in the embryo and is not detectable in the endosperm. It shows also the clear pattern of an axis polarity, being more intensive in the areas of leaves and root primordia.

In summary, *ZmFIE2* gene is expressed in the embryo sac before pollination and in developing embryo after fertilization, as well as in vegetative tissues. This pattern of expression is very similar to that observed for *Arabidopsis FIE*, but very different from that observed for *ZmFIE1*.

Pattern of Maternal and Paternal ZmFie Allele Expression During Kernel Development.

The Arabidopsis *FIE* gene demonstrates a parent-of-origin effect on seed development, suggesting that only the maternal *FIE* allele is essential, whereas the paternal *FIE* allele plays no role in seed development (Yadegari *et al.* (2000) *Plant Cell* 12:2367-2382; Luo *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97(19):10637-10642). Current evidence supports the model that the *FIE* gene is an imprinted gene, in which the maternal allele is expressed and the paternal allele is silenced during seed development (Yadegari *et al.* (2000); Luo *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97(19):10637-10642). To understand whether the maize *FIE* homologues are regulated by imprinting in the same manner as the *Arabidopsis FIE* gene, the paternal- and maternal-specific *FIE* mRNA levels were measured in developing kernels.

To distinguish maternal and paternal *ZmFIE* mRNAs, the insertion/deletion sequencing polymorphism was identified in both *ZmFIE1* and *ZmFIE2* genes in inbred lines Mo17 and B73. Reciprocal crosses were performed between B73 and Mo17 lines, and kernels were collected at 2, 5, 10, 15, and 16 DAP. Ovules and selfed kernels from both inbred lines were sampled at 11 DAP as controls. Total RNA was extracted from the whole kernels.

Mo17 and B73 *ZmFIE1* alleles are different by a 12 nt insertion/deletion in the 3' UTR. The reverse and forward primers were designed around this indel to

produce the 300bp RT-PCR product, which was separated on D-HPLC column by WAVE machine. As shown in Figure 2A and 2B, only maternal *ZmFIE1* RNAs were detected in reciprocal crosses in 15 DAP kernels. No detectable level of the paternal RNA was found at early stages (data not shown). The same set of RNAs was used with an anonymous gene as a control for bi-allelic expression (Figure 2C). The paternal allele of a control non-imprinted gene was detected in 5 DAP kernels and all later stages, confirming that the paternal gene is expressed in kernels. Thus, the *ZmFIE1* paternal allele undergoes transcriptional silencing in developing kernels, and this gene is regulated by imprinting. As noted above, *ZmFIE1* is expressed predominately in endosperm; this is in agreement with previous reports that all known imprinted genes in plants are expressed in triploid endosperm. Thus far, imprinting has not demonstrated for genes expressed in diploid tissues.

A different strategy was used for monitoring allelic expression of *ZmFie2*. *ZmFie2* genomic sequence from inbred B73 contains the 185 nt MITE insertion at 3' UTR, which is not present in the Mo17 allele (Figure 3A). The insertion is flanked by 15-nt inverted repeats and creates the 5 nt direct target duplication (Figure 3B). These features are typical for MITE elements, which are very abundant components of the maize genome (Wessler (2001) *Plant Physiol.* 125(1):149-51). In B73, *ZmFIE2* polyA transcripts are terminated in the middle of the MITE insertion. In Mo17 background, *ZmFIE2* polyA transcripts are terminated within genomic sequence with no homology to MITE. The MITE sequence was used to design allele-specific primers to discriminate between B73 and Mo17 *ZmFIE2* mRNAs (Figure 3A).

The forward primer, F, designed for exon 11, is common for both genotypes. The reverse primers, R, are genotype specific. The primer combinations are highly allele- specific; no RT-PCR products are found in RNA samples from ovules or selfed homozygous kernels. The primers allow monitoring of the expression of maternal and paternal *ZmFIE2* alleles in developing kernels. Maternal allele expression was detected at all stages in both reciprocal crosses, being more abundant in 2 DAP zygotes. These results are in agreement with the *in situ* hybridization data, which demonstrated an increased *ZmFIE2* expression in 2 DAP zygotes in the embryo-surrounding region. Paternal allele expression is delayed up to 10 DAP, but at later stages, both maternal and paternal alleles are expressed. Delayed expression, but not a complete silencing, of the paternal allele is a feature of the *Arabidopsis FIE* gene. As mentioned previously, the *ZmFIE1* gene undergoes permanent silencing of the paternal allele, demonstrating a different type of imprinting.

Genomic Structure of ZmFIE Loci.

The transcriptional pattern of *ZmFIE* genes is very different with respect to tissue specificity, efficiency, and imprinting. *ZmFIE1* is expressed only in developing kernels at a relatively high level, and with a permanent silencing of the paternal allele. Conversely, *ZmFIE2* is expressed in vegetative and reproductive tissues, showing a very low level of expression in developing kernels, with delayed paternal allele expression. To reveal the molecular mechanisms underlying the different patterns of *ZmFIE* expression, the genomic loci of both genes have been sequenced. Genomic BAC libraries were screened with *ZmFIE1* and *ZmFIE2* cDNAs. Five BACs were identified for each gene covering the overlapping regions (about 250 kb). Approximately 12-kb segments carrying *ZmFIE* genes have been subcloned and sequenced (Figure 4A and 4B) . The positions of nucleotides are relative to the translation start site, ATG (+1). (The transcription start site is used more often as a reference point, but it is not identified precisely for FIE transcripts.)

The coding regions of both genes downstream of the translation start site, ATG, possess 13 exons, which are identical in size between the two genes, except for the first and last exons where initiation and termination of transcription occur. The number and sizes of the protein coding exons are also identical to the *Arabidopsis FIE* gene (GenBank Accession No. AF129516). The intron sequences vary in length and do not share a significant homology between *ZmFIE1* and *ZmFIE2* and *Arabidopsis*. However, *ZmFIE1* demonstrates a unique feature among the FIE family, the presence of a 290 bp intron, located in the 5' UTR, just 6 nucleotides upstream from the ATG codon (-6 and 390). The first exon and intron are very often required for high level expression of the reporter, which may be a result of the increased level or stability of the mature cytoplasmic mRNA constructs (Kim and Guiltinan (1999) *Plant Physiol.* 121(1):225-236); Clancy *et al.* (1994). It is very likely that the 5' UTR intron of *ZmFIE1* plays a regulatory role or determines the tissue specificity of FIE1 protein expression.

The 5' upstream regions of the two genes are very different. The size of the putative promoter region of the *ZmFIE1* gene is estimated to be about 900 nt, between the RNA start of the longest EST (Accession No. AY061964) and the retrotransposon *RIRE* LTR (Figure 4A; Figure 5). Dot plot analysis (data not shown) does not reveal any repeats as far as 5 kb upstream of the *RIRE* retrotransposon. Repeats are commonly speculated to be involved in imprinting (Alleman and Doctor (2000) *Plant Mol. Biol.* 43:147-161). However, this analysis indicates that this is very unlikely to be the case for the imprinting mechanism of the *ZmFIE1* gene.

The 5' upstream region of the *FIE2* gene is about 6 kb long as estimated between the transcription start site of the *ZmFIE2* longest cDNA (Accession No. AY061965) and the retrotransposon *MILT* LTR. The extensive BLAST search of this sequence against the public and proprietary databases did not show any homology to known sequences, suggesting that the 6 kb 5' upstream region of the *ZmFIE2* gene is its unique integral part. Dot plot analysis (not shown) revealed the complex pattern of repeats positioned along the 6 kb upstream region (Figure 4B; Figure 6). The sequence between -1161 and -3479 consists of three types of repeats, named A, B, and C. Repeats form a 2.6 kb symmetrical structure having the following order: A1-B1-C1-B2-A2. The B3 and C2 types are repeated again (-5328 to -6077) forming one more cluster. Repeats A1-A2 are 550 nt long with 95% homology; B1-B2-B3 are 350 nt long with 94% homology, and C1-C2 are 420 nt long with 93% homology (Figure 6). Repeats do not share any homology or features of the transposable elements. They form a unique configuration and may be considered as a potential cis-regulating element of the *ZmFIE2* gene. The basal promoter of the *ZmFIE2* gene is estimated to be about 768 bp if framed between -393 and -1161, which marks the transcription start of the longest EST and the beginning of the B2 repeat.

The CG Composition of the ZmFIE Genes in Relation to Imprinting.

As discussed above, *ZmFIE* expression is regulated by imprinting but in a different temporal fashion. The paternally derived *ZmFIE1* allele is permanently silenced during kernel development. Expression of *ZmFIE2* undergoes less stringent temporal imprinting, because the paternal allele is reactivated later in kernel development (after 10 DAP). It has been widely speculated that imprinting is mediated by DNA methylation. CpG island methylation may be a key molecular mechanism of imprinting (Wutz *et al.* (1997) *Nature* 389(6652):745-749; Thorvaldsen *et al.* (1998) *Genes Dev.* 12(23):3693-3643; Reik and Dean (2001) *Electrophoresis* 22(14):2838-2843). Recently a two-island rule was proposed to define genes regulated by imprinting (Onyango *et al.* (2000) *Genome Res.* 10(11):1697-1710). In this reference, comparative analysis of human and mouse imprinted genes revealed that two or more CpG islands are associated with imprinted genes, while at most one GpG island is associated with nonimprinted genes. The CpG islands were defined in this reference as sequences of about 200 bp with a GC content >50% and an observed-to-expected CpG content >60%. These criteria were applied for searching for CpG islands along the *FIE* loci.

This analysis revealed three CpG islands within the *ZmFIE1* locus. One island is located between -2968 and -3219 (Figure 7), which corresponds to the retrotransposon segment and very likely is irrelevant to regulation of *ZmFIE1*. The

other two islands are located within the *ZmFIE1* coding region, which agrees with the two-island rule. The first of these two CpG islands is 252 bp and is positioned between +87 and +374, just downstream of the ATG codon. The second of these two CpG islands is 572 bp long and is located at the 3' end of the gene, between +4315 and +4886, covering the last two introns and exons.

Only one CpG island is present in the *ZmFie2* locus, at position -231 to +88, around the ATG codon (Figure 7). This agrees with the definition of non-imprinted genes, which are associated with at most one CpG island (Onyango *et al.* (2000) *Genome Res.* 10(11):1697-1710).

These data suggest that the imprinting mechanism of *ZmFIE1* is very likely associated with DNA methylation of two CpG islands. The delayed expression of the paternal *ZmFIE2* allele, which could be considered as a temporal imprinting, is not associated with DNA methylation. The complex repetitive structure of the 5' upstream region may be responsible for this type of imprinting.

Phylogenetic Analysis of Plant FIE Proteins.

ZmFIE1 and *ZmFIE2* genes are mapped to chromosome 4 (bin 4.05) and chromosome 10 (bin 10.3). These regions are duplicated in the maize genome (Helentjaris (1995) *Maize Newsletter* 69:67-81; Gaut and Doebley (1997) *Proc. Natl. Acad. Sci USA* 94(13):6809-6814). It is very likely that the two *ZmFIE* genes are due to the allotetraploid origin of the maize genome (Gaut and Doebley (1997), *supra*. Presence of two *FIE* genes in the maize genome raises the question whether two *FIE* genes exist in other species as well. A search by TBLASTX of the public EST database reveals accession numbers for 11 species, and putative FIE proteins were reconstructed. The FIE protein belongs to the *Polycomb* Group (PcG) proteins, which include the *Drosophila* extra sex combs (ESC), and mammalian embryonic ectoderm development proteins (EED). To make the phylogenetic analysis more robust, the PcG proteins from five insect and two mammalian species were included. A phylogenetic tree was constructed using the PAUP program (Figure 8). The phylogenetic tree forms four major clades corresponding to mammals, insects, monocots, and dicots. The *Arabidopsis* FIE protein is positioned apart, reflecting the absence of the protein from related species.

So far all analyzed plant species show the presence of the one putative FIE protein. The phylogenetic tree demonstrates that the sorghum FIE and ZmFIE2 proteins are more closely related to each other than *ZmFie1* protein. Thus, a *ZmFie1* analog has not yet been found. This does not prove the absence of homologs to ZmFIE1 in other species, but the probability is very high that FIE1 is unique to the maize genome.

Discussion

ZmFIE Genes Are Differentially Expressed.

In understanding the role of *ZmFIE* genes, it is crucial to know in which tissues and cells these loci are active and whether two genes are active in the tissues at the same developmental times. The *Arabidopsis* single FIE gene (*AtFIE*) is expressed in many tissues, both reproductive and vegetative, indicating that this FIE protein may have multiple functions during plant development. *AtFIE* is expressed in the embryo sac before fertilization, and its expression continues in the embryo and endosperm after fertilization (Ohad *et al.* (1999); Luo *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97(19):10637-10642) Loss-of-function alleles of *AtFIE* demonstrate pleiotropic phenotypes, including initiation of endosperm development without fertilization, embryo abortion at early stages, premature flowering by seedling shoots, and flower-like structures along the roots and hypocotyls (Ohad *et al.* (1999) *Plant Cell* 11:407-416; Kinoshita *et al.* (2001) *Proc. Natl. Acad. Sci. USA* 98(24):14156-14161). These results suggest FIE protein encoded by a single-copy gene in the *Arabidopsis* genome may form distinct complexes in different plant tissues and participate in repression of several developmental programs.

As has been shown by RT-PCR, the *ZmFIE1* gene is active only in kernels after pollination, but *ZmFIE2* has very broad expression in virtually all tissues, much like the *Arabidopsis FIE*. Because both *ZmFIE* genes are expressed in developing kernels, their expression in this organ have been studied by different methods to understand whether these genes have a functional redundancy. The RNA gel blot experiments revealed significant differences between the transcriptional activity of these two genes. *ZmFIE1* RNA revealed the inducible pattern of expression with a maximum activity around 9 DAP. *ZmFIE2* RNA is detected at a steady level across the various developmental stages as very low-abundance transcripts. Moreover, the *FIE* genes are active in different tissues of the developing kernels. *ZmFIE1* is active in the endosperm, as shown by the MPSS RNA profiling experiments (Figure 1). The small number of tag sequences detected in the embryo tissues may be explained by contamination of the embryos with endosperm cells during tissue dissection, particularly in view of the sensitivity of detection in MPSS experiments (1 molecule per million). *ZmFIE2* cDNA is not suitable for MPSS analysis, as it lacks the restriction site for enzyme DpnII, which is used to generate tags. But, *in situ* hybridization experiments have shown that the *ZmFIE2* transcripts occur in the embryo, not in the endosperm, suggesting that these two *ZmFIE* genes are active in different tissues of the developing kernels. Thus the expression patterns argue in favor of the nonredundant function of these two FIE proteins in developing kernels.

Of importance is the pattern of *ZmFIE* expression in the female gametophyte, i.e., the embryo sac before fertilization. The *Arabidopsis FIE* mRNA is found before fertilization in the embryo sac (Luo *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97(19):10637-10642), confirming its function as a repressor of endosperm development. Expression of *ZmFIE* is different in the female gametophyte as well. *ZmFIE1* mRNA is not detected in ovules by RNA blot analysis or by MPSS profiling (Figure 1). The high sensitivity of MPSS provides strong evidence of no basal expression, or low expression, of *ZmFIE1* in ovules before pollination. Conversely the *in situ* hybridization data show a detectable amount of *ZmFIE2* RNA in the embryo sac. Out of these two maize FIE proteins, only FIE2 is a candidate for a repressor of endosperm development before fertilization, the function performed by the *Arabidopsis* FIE protein. Loss-of-function mutant analysis will confirm this function.

ZmFIE Genes Are Regulated by Imprinting.

The prominent feature of the *Arabidopsis FIS* genes is their parent-of-origin effect in developing seeds (Grossniklaus *et al.* (1998) *Science* 280(5362):446-450; Ohad *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93(11):5319-5324; Luo *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97(19):10637-10642). The wild-type paternal alleles do not rescue the maternally derived mutant alleles (Grossniklaus *et al.*(1998) *Science* 280(5362):446-450; Ohad *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93(11):5319-5324); and paternally derived allele expression is delayed (*FIE* and *MEA*) or nonexistent (*FIS2*) (Luo *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97(19):10637-10642). *FIS* genes are regulated by imprinting, emphasizing the importance of maternal control of early seed development.

To investigate the possibility that *ZmFIE* genes are also imprinted, several experiments were conducted to monitor the paternal and maternal *FIE* RNAs in developing kernels. Both genes show silencing of paternal allele expression with a distinct temporal pattern.

The *ZmFIE2* paternal allele shows no detectable activity until 10 DAP. This pattern of silencing is very similar to *AtFIE* in which imprinting is in force until 3 DAP and later breaks down (Luo *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97(19):10637-10642).

The *ZmFIE1* paternal allele shows no expression at any developmental stages (Figure 2), resembling in this aspect the *Arabidopsis* gene *FIS2* (Luo *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 97(19):10637-10642). *ZmFIE* and *AtFIS2* are different types of proteins, but they are encoded by genes with very specific patterns of expression in the endosperm. *FIS2::GUS* activity was observed only in endosperm of the developing seed (Luo *et al.* (2000) *Proc. Natl. Acad. Sci. USA*

97(19):10637-10642). *ZmFIE1* expression is also limited to the endosperm. This suggests that genes that are expressed only in endosperm, similar to *AtFIS2* and *ZmFIE1*, undergo more stringent, permanent imprinting. Genes that are expressed in both embryo and endosperm, like *AtFIE* and *MEA*, are regulated by less stringent, temporal imprinting, which causes a delay in expression of paternal alleles, and subsequent breakdown of imprinting later in development. The *ZmFIE2* gene belongs to this group, which is regulated by a temporal type of imprinting.

Mammalian Models of Imprinting May Be Applicable to Plants.

ZmFIE genes have a differential parent-of-origin activity and are regulated by permanent and temporal types of imprinting. The presence of repeated sequences is a common feature of epigenetically silenced and imprinted genes (Alleman and Doctor (2000) *Plant Mol. Biol.* 43:147-161). Fragments of 12 kb of the Mo17 genomic loci of *ZmFIE* have been sequenced (Figure 4). A complex repetitive structure is found 5' upstream of the *ZmFIE2* coding region. Repeats occupy the 2.6 kb fragment adjacent to a putative promoter and a 1 kb fragment further upstream. The entire 6 kb upstream fragment does not share any homology to transposable elements, which are abundant sequences of the intergenic regions in the maize genome. It appears that the structural repetitive complex upstream of the *ZmFIE2* gene is an integral part of this gene and may be a cis-element regulating *ZmFIE2* activity. A critical aspect of the *ZmFIE2* expression is the delayed activity of the paternal allele in the developing kernels, referenced herein as temporal imprinting. The upstream-positioned repeats may be involved in setting imprinting marks on the *ZmFIE2* gene during gametogenesis. It is possible that specific proteins that function as activators or repressors of gene expression bind with these repeats. These complexes might be temporally associated with the upstream sequence but degraded during kernel development.

The genomic sequence of the *ZmFIE1* gene does not possess such obvious structures as repeats. Moreover, the promoter region of *ZmFIE1* is relatively short, approximately 780 nt between the putative RNA start and the LTR of a retrotransposon RIRE (Figure 4). The special feature of the *ZmFie1* gene is the 290 bp intron positioned at the 5' untranslated region. The first exon and intron are often required for high level expression of the reporter that may be a result of the increased level or stability of the mature cytoplasmic mRNA constructs (Kim and Guiltinan (1999) *Plant Physiol.* 121(1):225-236); Clancy *et al.* (1994)). It is very likely that the 5' UTR intron of *ZmFIE1* plays a regulatory role or determines the tissue specificity of FIE1 protein expression. There are no indications in the

literature that introns are involved in genomic imprinting. It has been proposed that CpG islands might be common imprinting elements in mammalian genes regulated by imprinting (Wutz *et al.* (1997) *Nature* 389(6652):745-749). Methylation of these islands during gametogenesis create the imprinting signals that maintain expression of the maternal or paternal alleles. The comparative analysis of mouse and human imprinted domains suggests a two-island rule for imprinted genes (Onyango *et al.* (2000) *Genome Res.* 10(11):1697-1710). Imprinted genes show two or more conserved CpG islands upstream or with the gene, while non-imprinted genes are associated with at most one CpG island. CpG islands are normally unmethylated and associated with actively transcribed genes, but allele-specific methylation of CpG islands appears to mark imprinted genes in mammals (Wutz *et al.* (1997) *Nature* 389(6652):745-749).

The distribution of CpG islands within the *ZmFie1* and *ZmFie2* genomic sequences was searched using a definition of CpG islands as sequences of >200 bp with a GC content >.5 and an observed-to-expected CpG dinucleotide content >0.6. This analysis revealed two CpG islands in *ZmFIE1* and one CpG in *ZmFIE2* (Figure 7). The results concur with a two-island rule. The *ZmFIE1* gene, in which the paternal allele is silenced during all stages of kernel development, shows two CpG islands. The *ZmFIE2* gene, which demonstrates a more relaxed type of imprinting, shows only one CpG island, implying a different mechanism of delayed expression of the paternal allele, which is not associated with DNA methylation. The data presented herein suggest that CpG islands may be the imprint marks in plants as well.

This assumption generates several predictions that may be experimentally tested. Transgenic constructs with a reporter gene placed between CpG islands should mimic the parent-of-origin pattern of expression of the *ZmFIE1* gene. A pattern of DNA methylation across the *ZmFIE1* gene can be tested in DNAs isolated from the male and female gametophytic tissues (pollen and ovules), and endosperm. This would provide evidence for differential methylation of the islands during gametogenesis and its maintenance during endosperm development. Further, imprinted antisense transcripts are observed in all major imprinting models in mammals (Fu *et al.* (2000) *Proc. Natl. Acad. Sci. USA* 99(2):1082-1087), which were proposed originally as the sense/antisense competition model for preferential allelic expression of the mouse *Igf2r* gene (Wutz *et al.* (1997) *Nature* 389(6652):745-749).

The two-island rule can be used to predict imprinted genes in plants. In this manner, a search of 2,000 full-length transcripts of annotated genes reveals that 10% of them fall within the category of two and more CpG islands. Relatively few genes are described in plants as being regulated by imprinting, but this approach

provides a potentially useful predictive tool for identification of imprinted genes. Support for the relevance of this approach comes from the finding of the α tubulin cDNA (tubα4), which shows two CpG islands. Imprinting of the maize α tubulin genes (families tubα3 and tubα4) has been documented (Lund *et al.* (1995) *Mol. Gen. Genet.* 246(6):716-722). Moreover, expression of the sense and antisense transcripts of the α tubulin genes were demonstrated earlier (Dolfini *et al.* (1993) *Mol. Gen. Genet.* 241(1-2):161-169). Having demonstrated the applicability of the two CpG island rule for imprinting in the maize FIE genes, it seems probable that this rule operates generally in plants, and suggests that the general mechanism of imprinting may be conserved in evolution across the kingdoms.

Two FIE Genes Reflects the Maize Genome Evolution.

The *ZmFIE* genes are located in the regions of chromosome 4 and chromosome 10, which are very likely duplicated in the maize genome (Helentjaris (1995) *Maize Newsletter* 69:67-81; Gaut and Doebley (1997) *Proc. Natl. Acad. Sci. USA* 94(13):6809-6814). The phylogenetic analysis of the known plant FIE proteins shows that sorghum and the maize FIE2 protein are more closely related to each other than to the maize FIE1 protein (Figure 8). This observation concurs with the hypothesis that the maize genome is a product of a segmental allotetraploid event (Gaut and Doebley (1997) *Proc. Natl. Acad. Sci. USA* 94(13):6809-6814). These authors provided evidence that "at least some elements of the sorghum genome share a more recent ancestor with one of the two maize subgenomes than the two maize subgenomes share to each other" (Gaut and Doebley (1997) *Proc. Natl. Acad. Sci. USA* 94(13):6809-6814). One can speculate that a segmental duplication of chromosome 10 around a centromeric region (Bin 10.03) has its origin from the sorghum-related progenitor. The orthologous region on chromosome 4 around the centromeric region (bin 4.05) carrying the *ZmFIE1* gene might originate from the second ancient genome that was more diverged from sorghum.

Despite the similarity between *ZmFIE1* and *ZmFIE2* genes, they are differently regulated. The *ZmFIE2* gene has a broad expression pattern whereas *ZmFIE1* expression appears to be restricted to developing kernels. These genes are regulated by different types of imprinting. The data herein strongly support the nonredundant function of these genes. *ZmFIE2* gene is very likely to be a functional homologue of the *Arabidopsis* FIE genes with multiple functions during maize development, such as preventing endosperm development before fertilization, and may be involved in functions for embryo growth and control of flowering. The second maize gene, *ZmFIE1*, has evolved for a kernel-specific

function, most likely in endosperm development. Experiments with null mutant analysis will further elucidate the function of these genes in maize.

Example 2: Imprinting of the Maize Endosperm-Specific Gene FIE1 Is Mediated by Demethylation of the Maternal Complements

Significant progress has been made on revealing imprinting mechanisms in mammals, but no such progress has been made in plants. The underlying mechanism of mammalian imprinting is differential DNA methylation of maternal versus paternal alleles, a process that takes place during gametogenesis (Constancia M, et. al., Genome Res. 1998, 8: 881-900). DNA methylation means the occurrence of 5-methylcytosine instead of cytosine in the context of CpG sequence. The major function of cytosine methylation is transcriptional repression.

Most of the CpG sites in higher eukaryotes are methylated with the exception of CpG islands, which are stretches of DNA enriched in CG di-nucleotides (Ponger et. al., 2001,Genome Res 11: 1854-1860). Imprinted mammalian genes show differential DNA methylation in CpG islands (Reik, et al.,2001, Nat Rev Genet 2, 21-32). Onyango et al. (Genome Res, 2000, 10:1697-1710) reported that the mammalian imprinted genes show two or more CpG islands within gene sequences, an observation referred to as the two-island rule. As shown herein, the maize FIE1 gene is imprinted and contains two CpG islands in its genomic sequence. This suggests some similarity between imprinting mechanisms in plants and mammals. The role of cytosine methylation in imprinting of the ZmFIE1 gene was further investigated, as follows.

Results

DNA methylation assay of ZmFIE genes in leaves, embryos and endosperms. To investigate whether cytosine methylation occurs within ZmFIE genes and correlates with imprinting, a quick and simple method was developed; it comprises DNA digestion with methylation-sensitive restriction enzymes, followed by PCR amplification across the restriction sites. PCR amplification of digested DNA occurs only if the cytosines were methylated and thus protected the DNA from digestion.

Commonly used enzymes *HpaII* and *MspI* were chosen for this analysis, but any other methylation-sensitive enzymes or mixture of several enzymes could be used. Both enzymes recognize CCGG sites, but show different sensitivities to cytosine methylation (New England Biolab catalog). HpaII does not cut DNA if either cytosine is methylated. MspI cuts DNA with the internal cytosine methylated, but does not cut DNA when the external cytosine is methylated.

PCR primers positioned across the restriction CCGG sites will amplify the HpaII/MspI digested DNA if CCGG sites are methylated. PCR reaction on unmethylated HpaII/MspI digested DNA will fail.

The restriction maps of ZmFie1 and ZmFie2 genomic sequences (Figure 9) show a distinct distribution of HpaII/MspI sites (CCGG) across the genes, scattered along ZmFIE1 and grouped in a cluster in ZmFIE2.

As shown previously, the ZmFIE1 gene has two GC-rich segments defined as CpG islands. The first island is located within exon 1. The second island covers exons 11-12 and 3'UTR. The islands have two and three HpaII sites, respectively. There are also HpaII sites in exon 7 and exon 10. Four pairs of primers were designed around clusters of HpaII sites to monitor cytosine methylation in CCGG sites (Figure 10).

The ZmFIE2 gene has one CpG island within exon 1. Eight HpaII sites are grouped there. No HpaII sites are present in any other segments of the ZmFIE2 gene. One pair of primers was designed for the ZmFIE2 gene (Figure 10).

DNA samples isolated from embryos and endosperms of 14DAP kernels of reciprocal crosses between public inbred lines B73 and Mo17 were digested with HpaII and MspI enzymes separately. DNA extracted from leaves of B73 inbred was used as a control. PCR amplification of an equal amount of undigested and digested DNA was performed and PCR products were visualized on agarose gels.

For the ZmFIE2 gene, none of the digested DNAs support PCR amplification, indicating that CCGG sites within ZmFIE2 are unmethylated in tissues tested (leaves, embryos, endosperms). These results are in good agreement with the expression pattern of the ZmFIE2 gene. As shown previously, this gene is expressed in all tissues throughout development. The unmethylated status of the gene is consistent with its transcriptional activity.

Conversely, a specific pattern of cytosine methylation across the ZmFIE1 gene was found. CCGG sites within CpG island 1 and exon 7 are methylated in both cytosines because HpaII and MspI digested DNAs are amplified effectively. This pattern of cytosine methylation is present in all tissues tested (leaves, embryos, endosperms). But CpG island 2, which is located in the downstream portion of the gene, is methylated very weakly in embryo and leaf DNA, and is barely detectable by PCR in the endosperm. Results clearly demonstrate that there is a gradient of cytosine methylation along the ZmFIE1 gene, being heavily methylated at the 5' end and unmethylated at the 3' end of the gene. DNA methylation of the ZmFIE1 gene correlates well with a repressed status of this gene in all maize tissues except the endosperm. As was shown previously, only maternally transmitted ZmFie1 allele is expressed in the endosperm; maternally transmitted ZmFie1 allele must be demethylated in the endosperm DNA.

Maternally derived fie1 alleles are demethylated in the endosperm

Status of cytosine methylation of the maternally- and paternally- transmitted ZmFIE1 alleles in the endosperm DNA was determined by means of two SNPs (single nucleotide polymorphism) present in exon 1 of B73 and Mo17 inbred lines (Figure 11). PCR primers were designed around the SNPs and HpaII sites. If both alleles were methylated at $^{m}C^{m}CGG$ sites, the sequences of PCR products would show traces of both SNPs. If only one allele were methylated at $^{m}C^{m}CGG$ sites, the sequence of PCR products would have SNPs from only one parent.

To facilitate direct sequencing of PCR products, ZmFIE1gene-specific primers were extended with T3 and T7 primers at 5'ends. DNA isolated from embryos and endosperms of the B73 and Mo17 reciprocal crosses was digested to completion with HpaII and MspI enzymes. The digested DNA was amplified by PCR, and the fragments were sequenced with T3 and T7 primers. Chromatograms of the nucleotide traces of PCR products from embryo DNAs showed a mixture of SNPs from both parents, B73 and Mo17. This is strong evidence that both parental alleles are methylated in the embryo. Conversely, the chromatograms of PCR products generated from the endosperm DNA show SNPs from the paternally transmitted alleles and complete absence of traces from the maternally transmitted alleles. Undigested DNAs, used as a control, showed a mixture of traces from both parents.

Discussion

This indicates that the ZmFie1 paternal allele remains methylated in the endosperm, but the maternal allele undergoes de-methylation followed by transcriptional activation. Data suggest that the methylated state is a default for the FIE1 gene; thus transcriptional activation of the maternal fie1 complements is achieved through demethylation. The paternal allele remains methylated and transcriptionally inactive during endosperm development. Maternal-specific de-methylation explains the mechanism of imprinting of the ZmFIE1 gene. It is very likely that demethylation of the maternal genes is taking place in the central cell of the female gametophytes before fertilization.

All publications and patent applications mentioned in the specification are indicative of the level of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be

obvious that certain changes and modifications may be practiced within the scope of the invention.

WE CLAIM:

1. A method of identifying imprinted genes in a plant, comprising identification of two or more CpG islands located partially or completely within the coding region.

2. The method of Claim 1 wherein said plant is of the species *Zea mays*.

3. A method of identifying plant genes involved in endosperm development, comprising identification of two or more CpG islands located partially or completely within the coding region of said genes.

4. A method of silencing paternally-transmitted alleles of a plant gene, comprising transformation of a plant with a construct comprising at least two CpG islands operably linked to the coding sequence of the gene of interest and a promoter that drives expression in plants.

5. A method of detecting cytosine methylation in a polynucleotide of interest, comprising:

(a) Restriction of said polynucleotide with methylation-sensitive restriction enzymes, followed by

(b) PCR amplification using primers positioned across the restriction sites for the methylation-sensitive enzymes wherein PCR amplification of digested DNA will occur only where methylation protects the polynucleotide from restriction.

6. A method of controlling plant gene expression in the endosperm, comprising demethylation of CpG islands in the allele contributed by the female parent.

7. The method of Claim 6, wherein the plant is of the species *Zea mays*.

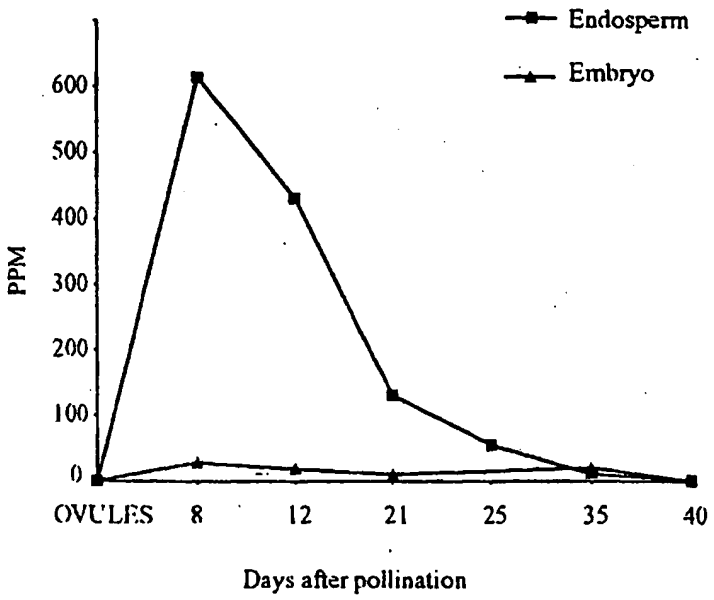

Endosperm
Embryo
600
500
400
300
200
100
0
PPM
OVULES
8
12
21
25
35
40
Days after pollination

FIGURE 1

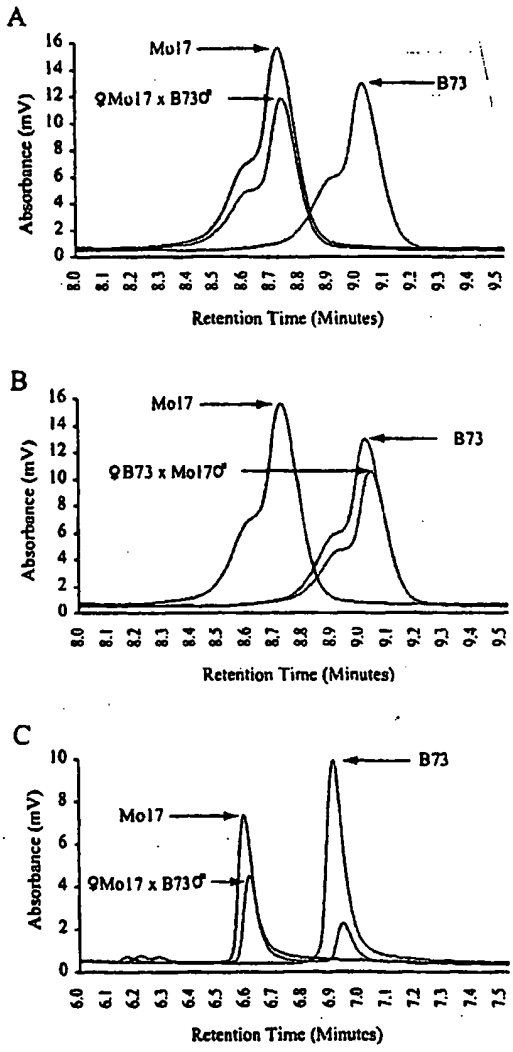
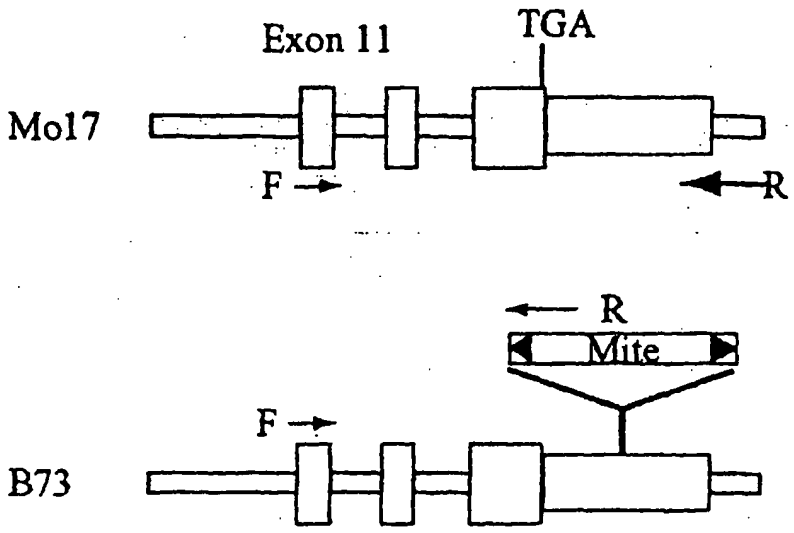

A
Mo17
B73
♀Mo17 x B73♂
Absorbance (mV)
Retention Time (Minutes)
B
Mo17
B73
♀B73 x Mo17♂
Absorbance (mV)
Retention Time (Minutes)
C
B73
Mo17
♀Mo17 x B73♂
Absorbance (mV)
Retention Time (Minutes)

FIGURE 2

A



B

CTTAGGTGCCGTTTGGTTCACATATTTGTAACGTAATGGGTAACAGATA
ACGTTAAATCATGTTTGTTTTATTTCAACCGTAATCAGATACCACATTA
AAATTTGATACCAGACTATTCAAATTTGTTAACGCCAGTAATCGAGCGC
AAACCATTACCATTTGCGTTACATTTTTTGAACCAAACAGCACCTTAG

FIGURE 3

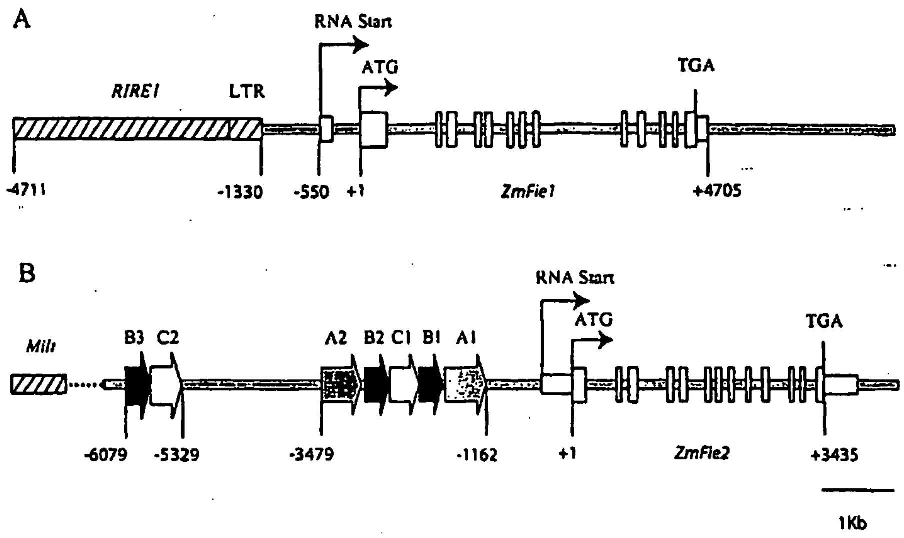

A
RNA Start
ATG
TGA
RIRE1
LTR
-4711
-1330
-550
+1
ZmFie1
+4705
B
RNA Start
ATG
TGA
Mil
B3 C2
A2 B2 C1 B1 A1
-6079
-5329
-3479
-1162
+1
ZmFie2
+3435
1Kb

FIGURE 4

>ZmFIE-1 5' upstream

SQ Sequence 5506 BP; 1538 A; 1215 C; 1048 G; 1703 T; 2 other;

CCGATCATTC GTTTGTTCGA TCATTTGATC GTTCATCGTT CGTTCATAGT TCCTATTCAT	60
CGTTCATCGT TTGTTCATAG TACTTATTCA TCGTTCATCG TTCGTTCATA GTTCCTATTC	120
ATCGTTCATC GTTACTATTC ATCGACACTA TTCACCATCG TTACTATTCA TTGTTACTAT	180
TTACCGGCTC TATTCGTCAT CGTTACTATT CATCGTTGCT ATTTATGGTA GCTTTTTCGT	240
TGTTACTATT CATCGATCAT CCGATCGCCC CAAATTTCAA CTACTCATCC ATCATGTTGT	300
CCAGTCCACC TAAGACCAGC CAGACCCATA TTCCAGTCAT ACGAACTCCT GTGATTGTGA	360
TTTTCCTTCC AGTAGGGAAC CTCCCATCTG GTCACCCATC CTAGGTTTCT CCAAGTTGAG	420
CATGCTTAAC TTTGAGATTC CTTTGAACCA GGCTTCCAAA CTCAGATTCC AATAATTCTT	480
GTTTCTAAAT TCTTATCAAA CTATTCCCTA TCCAACCATG TCATCCCTTA AGCCTGGTCC	540
ATATTCCAGA AAACTCCCAA AATACTCTTG TCCCATATTC TGCATATAAC TCTCCTGTTC	600
ATACTAAGTC AGACGATTCA TTCGTCACTA TTCTCACCAA CAGTGAACTT CACTGTGCTA	660
CACCACATAC ACTCAGCTAT AAATACACCC AGCTACCCTC TCCCTCTCCA CACACACTCA	720
ACACCCTCAG CCAAGGCAAA CACCTCACCC ACTCAGTTAC TCCGCTCTAC CGGCTACACG	780
CATAGTGTCG CTTCGCCTCC AGTCCACCCT CCTGGTAAGC ACCTCCGCTC CACCACCAGT	840
AATATCACAA CACCACATGA CACAGATTCT ACTCAAGACT CTACCCATCC ATATATCGCT	900
ATTCTGACCA CTATACTAAA TATTTGTTGG TATACTTGCT GGTTTGTATG TTTGCTTGTT	960
CATGTTGCAT AGTTATCGGA GCGTTCGTGC CATCACGTGG AGGCCAGATC TGCAAGTCTA	1020
CGCCAGGCGG TGGAGCCAGA AGCCAGTTCC GCGAGCTCTC CTTCCCCCTT CACTGGATAA	1080
GCACAGCAAG CTCACTGGAT CCCTTTGATG CATAAATTAC CTATGATTTT TCAACCACAA	1140
CCCTCAGCCT GTTATTTTAT GCATAATATG ATTTTGAGAC AAGTTATTAT GGCCACCCAG	1200
CCGCTTGTCG CAATCAATCC TTGATATATT TGTTACAAAT GATTTGAGAA AAGGTGTGAG	1260
TTTTCAAAAG AAAATGCTTT TCAAAATGTG TATGATGAAG GGTTTTCACC CTTATCACCT	1320
TTTAATAGGG ATGATCAAGG ACTCCCTGGT TTAGGGGAGG GCCTAAGGTG ATGGCTCAGC	1380
TGGTTTAGGT GTGAGCAGAA GGATTGTCCC CTCACATAAG GACCGATTTG TCATCCGTCA	1440
CTACCTGTAC TCATGATAAG TACAACCACT CGAGACTGTA TGGGCAATCA CTCAATCTGA	1500
ACTCGTACGG TCCAACCCTA GGGTTATGAA GGCTGGGGAG CACCGGGAGG ATAAGGAGGG	1560
AGAATGTTTT GTCCGGTTTG GACATGGCGG TGGCCTGACT CCTTCCGGTA TAACCGTTAA	1620
GGTAAGGACG TGCGAGGAAA GAAAGAGATC CGGCATTCGG GCCTCACGAC GGTGAGATCG	1680
CAGAAACCAG ACTAGTGGGT AAAGTGTACC CCTCTGCGCA GAGTTTGAAA ACCTATTCGA	1740
ATAGTCTGTG TCCACAGGAA TGGACGAGTC TGGTGTGGTA TGACAATTAG TGTTTTGTTT	1800
TCAAAAAAGA ATGTGCGTTT GAGAAAAGTG GTTTTTAAAA GGTCCGGCGG TTGAGCCGTG	1860
AGCTATGGTG GACGGGAAGT CCAGTAGCTG TTTTTGAAAA CGAAAACCAG TGGGAAACTG	1920
CTGAGATACC TGGATGGTTT AGTCCAGGGG ATTTTGTTCT AATATTGAAA AAAAATTCTT	1980
GCTCCTTTGG GAGAGGATGC GCTTTGCAAA ATACAAAATG TTTTACAAAA TAACCCTGCA	2040
TAAAATATTG TTGTTTCTGC AAAATATCCT GAGCTCCACA TATTCCATGC ATTATATCTG	2100
ATTTCCCCAT TCCGCGGGTG ATGGTGGGCT GCTGAGTACG TTTGTACTCA CCCTTGCTTA	2160
TTTGTTGTTT TTCAAAAAAA GGAGATCGGG TAAGAGTTAC GACTGTTCCC AACCTTGCCT	2220
GTGGTTGTTG GACCGCTGAT TTGCTTCGCT GCGTATATCG GGCTGCTTCA TCCCCACTCT	2280
GATGATATGT CCCAAGTTGT GGACCAACTC TTAAAGTTGA TCGCCACCTT TATAGGTTTG	2340
TCTCGTTTAA GCAGATCTGG AATCATTTGA TGTATAAATG TGTTTACTAG CCTCCTGGGA	2400
CTAGTAATTG TATCACATTT GAGTCCTAGA GGATCGGGAC GCTTCAATGA TCAATGGGTG	2460
GATCACAATA GTCGGTTATA ATGGCTATAT CAACAGTTAT AATCACATTA AATGTGTCAT	2520
CAGATGTTAG ATAAAGTCTG TCGTGGATGA TCTGTTTGTG CTTCTCGACG GTCCATGAGT	2580
GACGCTAAAA TTCATTTTAC CAAACCTAGC ACCTTCGAGT TGGTCTGATC TTGAATAGTC	2640
AGACGGTTCA CGACTGAGGT TGAACGATCC ACGCAAGGTG TTGGACGATA CTTTCTTTTT	2700
CTTTGGATGC TCCGTAGTAG ATGTGTCGGT TTTGACATAG TTCCTGTCCG AACTCCATAC	2760
AGTCCATAGT AGATGTGTCG GTTTTGGTAC TCTAGACGGC CCGAGTCAGG GGTCTGGACA	2820
GTCCTGGACT TGCTGAGTTG AGGTTTGATC TTTCTTTAGT TATTTCTTAC ATACCTATGT	2880
TCATACACTT AGCAAACTAG TTAGCTTCAC CAAAACAAGT GTGGAAAAAG GTTTTTAGGC	2940
CAATTTCCCT TTCACCTTTA TAACTACCTA GTTACAAAGT AGAGTTTGAT AGTCCCTAAG	3000
TATGTCAATT CACATCTTGA GTACATGCGA CAATCTCATG TCTAAGGATA CATGGTACAG	3060
GTTGCAAGAA GAAAATTGTC ACAATATCTC ATGTTGGGTC AGTACAGACT CATGTCATAC	3120
ATGCACCCAT ATTATTAGTT TTACATCTCC ATGTCCATGA CTTACGAAAC ATAGTCATCA	3180
ACTAATACAT ATGATAGTCA TTGACTCTAA CTAGGGACAT CTTCTAGAAC AACCATACAA	3240

FIGURE 5

```
GAAAAGAGTC TCACAAACAA TTCACATAAT TGCTAATCAA TACAAGGTGT CCTTCACAGA  3300
TATTCAATTA AACAATATAT CATGGATGCA ACAWAATATG CTCATCTCTA TGATTATCTC  3360
TAGGGCATAT TTCTAACACA ATGACATGTC TAAGTGTAGT ATGTCAAAAC ATGGATAGTA  3420
ATATAGATGG TAAGAGGTCA TTTTTATTAA TATAATTAAC AAAGATAGAT AGGGTGACCA  3480
ATTTTGTAAA AGCACCATTC ATAGACTTTT AGTGGGAGGT GGATGCTCTA CCCGCCTCCG  3540
TAAAGCCAAA GTGGTTGCAT GCAAATTGYT AGGATATAGT AATGCAAGGA ACCAAGCTAA  3600
GGCATGTAAG TGAAACCCAA ACAAGAAGTT AAGAAGCTTC CAAAATGAAC AAAGTACAAG  3660
AATGAAGCTA AAAGAGAAAC TTTCAGCCTT CTCCAATCTC CAGCAAGATC CCTTCGATAG  3720
ATGGTATCTA ATTTTTTCCT ACTATGAAAA CCTATATCAC CTAGTAGAAT AGAGGACAAA  3780
GCTTACGCCT ACTATATATA TCCAATATGT ATAGTTAGAT ACTAAGTTCT TTTTTCTCTT  3840
CTCTTCATTC ACTTTTCAAC TAGGTTTGGA ATTAAGTTTT TGGATTGGCA TAGACAATGG  3900
CATGGTTGTA TAGGTGTTCT TAACCATCAC AGTTATGAGT TTGACTTGTT TTTTATATTC  3960
AAGTTACAAG GTCATTTTGT GCTAGCCACA GCCTAGCAAT CGAGGGGCTA CACATGTGGA  4020
TTAAGGACAA GGCCCAACCC ATGTACGATC CAAGGACACC CTTGTAATTT TTATACTCAT  4080
CAAGGATTAG GGGGAAATAA CTCCCTTCTA TATAAAGGTC TTTCCACTTT GCTTCTCACT  4140
CTCCCTTATT AGGTTAAACA CAAAATGTGC ATCGCCGCCG CCACCATATA GAACCACTTA  4200
TCACGAACCG CCGCCATCAC ATCCACTGCC TCAACTAGTG TTACCACCTA TGGTTCATTG  4260
TTGTGTCTGC TTCTTGTAGC ACTGTTGGTC TACAAACATT CATATTTCTC TCAACATCTG  4320
GCACAGGTAA GCCCATAAGC CCTAACCCTA GATCTCCATA TTTAGTTATT TCAGTTCTTG  4380
ATGAGCAAAT ATGAAACTAA ATTAGTTTGC TAATAAGAAA TTTAACTACT TTTCCTCTTG  4440
AAGACCTCCT ATCCCTATAT GAACCCACAT CCAAAACCCC TCTAGCAAAG TGTGGCTAGC  4500
TTTCCCATGC CATGAACCTT CAACAATGAT AGTATCAGTA ATGCACTTCC ATAAAAGGGT  4560
TCATATTTAA TTTTAGTTTT TCTTTTTGGT GTTTTAATTA AGCTTTGAGA CTTGATTTGA  4620
AGTATTAAAT AAACCCTTCA AATTTCTTTC TAACTTTGAT AATACACTAT TCAATGACAA  4680
TGCACTTCCT TAAATCCCTA TACTTCACAG CATGCCGCCT TCCAAAGCAC GCCGAAAGAG  4740
GTCACTTCGT GATATCACTG CCACCGTTGC CACTGGGCCT GTTGCCAACT CGAAACCTGG  4800
CTCATCATCG ACGAACGAGG GGAAGCAACA TGACAAGAAA AAGGAGGGTC CACAGGAACC  4860
GGACATCCCA CCATTACCGC CGGTGGTGGT GAATATAGTC CCACGACAAG GATTAGGATG  4920
TGAAGTAGTG GAAGGGCTAC TCGTGCCTAG TCGGAAGCGA GAGTACAAGC CCAATAGCAA  4980
GTATACTGTG GGAAATCACC CGATCTATGC CATCGGGTTC AATTTCATTG ACATGCGCTA  5040
CTATGATGTC TTTGCCATCG CCAGTTGCAA TAGTGTAAGC AACCGACTTC TCCCTACCTC  5100
TTGTTTGCTA TCCTTTTATC CTATTGAGGT TTGGGGAGTT CTATATGGTG AACGAAAATG  5160
GAAGTTATGA TTTTGGTGGG ATTGGATCTT GGTTTATAAC TAGAAAAGGA TTTGAGTACA  5220
GGTTATGATG TGTGGCTTTA TGGTAGGGAA ACTTAATATC TTTTCCTATT TTGTTTTTTG  5280
GCATCACGAG TAATGGTTTG GGAAATAAAA GGGAAAATGA TTTAAAATTA TTTCTCAATA  5340
GAGCATGCCC TTTTACATAG GGACATTTTA GTCATTTTAC ACACACTTTA GTCATTTTAC  5400
ACACCGTAAT TATGTCACAA TCAAAGAATC ATTCCTTGGT TCAATTGAAT GAGATGATTC  5460
AACTAGTTCA CATCTCTATA CCTAACAATA TAGTTTTTCA TAACTA                  5506
>ZmFiel coding region
AAGCTTTGAG ACTTGATTTG AAGTATTAAA TAAACCCTTC AAATTTCTTT CTAACTTTGA    60
TAATACACTA TTCAATGACA ATGCACTTCC TTAAATCCCT ATACTTCACA GCATGCCGCC   120
TTCCAAAGCA CGCCGAAAGA GGTCACTTCG TGATATCACT GCCACCGTTG CCACTGGGCC   180
TGTTGCCAAC TCGAAACCTG GCTCATCATC GACGAACGAG GGGAAGCAAC ATGACAAGAA   240
AAAGGAGGGT CCACAGGAAC CGGACATCCC ACCATTACCG CCGGTGGTGG TGAATATAGT   300
CCCACGACAA GGATTAGGAT GTGAAGTAGT GGAAGGGCTA CTCGTGCCTA GTCGGAAGCG   360
AGAGTACAAG CCCAATAGCA AGTATACTGT GGGAAATCAC CCGATCTATG CCATCGGGTT   420
CAATTTCATT GACATGCGCT ACTATGATGT CTTTGCCATC GCCAGTTGCA ATAGTGTAAG   480
CAACCGACTT CTCCCTACCT CTTGTTTGCT ATCCATTTAT CCTATTGAGG TTTGGGGAGT   540
TCTATATGGT GAACGAAAAT GGAAGTTATG ATTTTGGTGG GATTGGATCT TGGTTTATAA   600
CTAGAAAAGG ATTTGAGTAC AGGTTATGAT GTGTGGCTTT ATGGTAGGGA AACTTAATAT   660
CTTTTCCTAT TTTGTTTTTT GGCATCACGA GTAATGGTTT GGGAAATAAA AGGGAAAATG   720
ATTTAAAATT ATTTCTCAAT AGAGCATGCC CTTTTACATA GGGACATTTT AGTCATTTTA   780
CACACACTTT AGTCATTTTA CACACCGTAA TTATGTCACA ATCAAAGAAT CATTCCTTGG   840
TTCAATTGAA TGAGATGATT CAACTAGTTC ACATCTCTAT ACCTAACAAT ATAGTTTTTC   900
ATAACTAGAA TTCTTAAAAA GAATTAATAT GAACCTAAAT ATTATTTCAC TTTCTTGCCC   960
CTTATAATAT AATACATTTG TCACTCCCAT TTTGGCAAGG GTGGTGGGTA TTTTGGGGGA  1020
TGGAATGTTA CTATTTTTAA TTTGATTAGA AGCTATAAGC TTTGGCTATA TTTTTATTAG  1080
```

FIGURE 5 (Con't.)

GAATTTGATG TTCATTTTCA ATATATTGTG ATCTATTTTC TTAAAATGTG AATTTGTTGT 1140
GTATTTTGAT TAGTTCGATG AAGAGTGTTT ATAAGATATG ATTTTTAAAT TCTCTTACGA 1200
CGAAACAATA TTATGTTACT TTCATCTATT CATCTTGAGG AATCACCTAC CTCACTTCTT 1260
GATCTTGCAG GTGATAATTT ACCGATGCCT TGAGAATGGT GGTTTTGGTC TTCTACAAAA 1320
TTATGTTGAT GAGGATGTGA GAAAGACAAT GCCTGGTGCA TGTGGTTGTT AATGTTAATT 1380
TGATAATATG CTTTTATCTA ATGTCTGTGG TGCCTATTTA TCTCAGAAGG ATGAGTCATT 1440
CTACACTCTA AGCTGGACCA TCGATCAAGT TGATAGCTCA CCGCTGTTGG TGGCCGCTGG 1500
AAGCAATCGG ATCATTCGGG TCATCAATTG TGCTACCGAA AAGTTAGATA AGGTCCCTGC 1560
CCCTGTGCTT ACTCTATGTT TGTATGGAAA AGTTGATTGA ACGTTGATGT TCACATATCA 1620
ATATTTCAGT AGTTTAGTTG AAATACAATT TATTTATGCT CTCTATTCTT GAACATCAGT 1680
TGACTTTGCT TTGATTAAGC AATGGTCTTG CTCATACAAT ATTCTAGGAG TTGAATATTC 1740
AATATGCCTG TTACATGATA GCAAATACAT AGTGAACTAG GACATGTACT AAATATTTAA 1800
TTTCCCTTTA TGACATTCTC TAGAGCTTAG TTGGCCATGG TGGTTCAATA CATGAGATAA 1860
GGACTCATGC CTCGAAGCCA TCACTCATCA TTTCTGCCAG CAAGGTTAGT AATAAATTTG 1920
TCGTGTGTCG ATTTTTTTAC ACTTTTTAAC ATGACATTAT TCTATAGGAT GAATCTATTA 1980
GGCTATGGAA TGTCCATACT GGGATTTGCA TCTTAGTCTT TGCAGGGGCT GGAGGCCATC 2040
GACATGATGT GTTGAGTGTT GTAAGTATCG ATTGCATCTT GTCTAGACAT TGTTTTAAAT 2100
ATCACTTGCC CCGAAGATAA CACTCATTAG AATTCTAATG TTACCATTTG TTATTGAGCA 2160
TGCCAAATTT CAATTTTAAC ATCATAGATA AAATAAGACC CCACAATTAC TTTTACTGTT 2220
TATCTACTTC CATTACATTA GGCATAAAGT TACTGATAAA AAAGACAATC TTTTATCTGA 2280
AGGACTTCCA CCCTACCGAG GTTGGGATTT TTGCAAGTTG TGGCATGGAC AATACTGTGA 2340
AGATTTGGTC AATGAAAGGT TTGGGAACTA CTTTAAACTA GCTTCATGTT TACATTTTGT 2400
GTTGTATGTT GCATATCATC GACAAATATT GCCAATGTTG TCACAGAATT TTGGATATAT 2460
GTTGAAAAAT CATATTCATG GACTGGCCAT CCATCAAAGT TTCCAACGAG GAATATCCAG 2520
TTTCCGGTAT GTTAAGTAGC TATAATCACC TGAGCTCCTT TCTTTTTTTG CAAACTATTG 2580
TTGGTGTTCA GTTTTCATGC CATTCAAGCA TACATGTTTC TTTTCTTTTA GGTCTTGACT 2640
GCTGCAGTAC ACTCTGACTA TGTTGATTGT ACCAAGATGG CTTGGTGACT TCATCCTATC 2700
AAAAGGTAAA TTCTTCATTT GTTAAATGGC TATACATTTT TTTATAAAGG AAATTTTTTA 2760
TTAATTTCAA GCACTTTAGA TTGAAATAAT ACAAAATCTT AAAAAACATT TTTGGCCTCC 2820
ATTTAAACAA GCACAAATCC AACAAAAATG AGTAAACCAA CCCATTCTAG TGAATATTAA 2880
TGCATAAACT AGATTGCTAC CCATATGTCT AGAAAAAGTA GCCTTGACCG CGTATCTTAA 2940
TTGTCACCAT GCCGCCACAA CCAAACCGTG CAAATATGGT TTTTGGAGAA TGGACCAAGT 3000
AAGAAACCAA TCAATAATTG AGTATATAGC ATGCACAGGA GAAATAGATC TCTTATTTTC 3060
AAGAACAATG GTATTTTTTA TTAACCATAG GACCAACAAG TAGCGACTAC CCATAGCAAA 3120
ACTAATGGCT TCAGATTATT ACTGGTTGTT GAAGTGTATA CGTGGTTTGC CTACTTTCTC 3180
CCAATAGTTT AAGCTTTTGG ATTGAATCGA TTAGTGCGTT CACTCTTACA TGGTATCAAA 3240
GTTAGCAATT TTGGGTTTGA ATCCTAACGG AAGCTTTATT TGTGACTTCA CCTCTTGTTT 3300
TCCATTTCCT TTCTACCTGC ACGTGAGTGG GGGTGTTGAA GTGTATAAGT GGATTGCCTA 3360
CCTTATCAAC CTTTTGGATT AAACTGGTTA TTGGTTAGTG TGTTCACTCC TACACCTAAG 3420
TATGAGGTTT AGTTATCCAG TAGCCAATTA GATTATGCAC AGTGGACACT TCACATGTGC 3480
AACTAGCACT CAAAACATAA GTCTTTAATT GTCTCATCTT ATGACAAAAC AACATATTTC 3540
ACTACCATTC TATAACATCT TGATTTGTAC ATCAGTCTTG TTAATGCTAA ATAGTGAGAT 3600
TTGATCGTCA ATTGGCCAGT TGGATGTAAA TTCCAGTGAA ATACATCTTG ACCTTGGGTT 3660
AAATGGACAT TAGCAATGTG TGGGAACAAA TTGTTGGTTT GGGTACACCA AACTGTTGGT 3720
TTTTAATTAG TAGATTAGTT TGTAACACAT TTCCTTTTAT CAGTGTTAGT ATTGGTTTAT 3780
TATGCATAGG GAAGGATCTG ATATGTGATA ATTAACATGG ATTTGCAGAG TGTAAAGAAT 3840
GCAGTTTTGC TTTGGGAACC AAAACCAGAC AAGCGTAGGC CTGGGGAGGT GACACGCTTT 3900
ACCTTCTCGT CCCGAATTCT GCACCTATTT TTATATTACT ATCATACTCA TCTACAGTTT 3960
AAAACTTGTC CCGCAATCTT TTCAGTTTCT GAGCACTAAA TTTATACCTC TGAATCAGTA 4020
TAGTCGTTTT CTCTTTGTTC GTATAGGGGA GTGTTGATGT TCTTCAGAAG TACCCGGTGC 4080
CAAAGTGTTC ATTATGGTTT ATGAAATTTT CATGTGATTT TTACTCCAAC CAGATGGCAA 4140
TAGGTAATGC CTTTAATTTT GTGAAGACTG TTTTGGCACT AAAGCTTTAC GTACGTAATA 4200
TTAGTTTTAT ATCTTGTACA TTGATGGAAA ATAGATTGCT CAATATCTAT ATATATGACT 4260
ATATCTTGGG TTAGATTCTA AGGAACAAAC TCTCCCAGAG TACGGTTCTG AATAACAACC 4320
ATCTGCTGCT GCTGCTTAAT GCGAACAGGC AACAATAAAG GCGAGATCTA TGTCTGGGAA 4380
GTGCAGTCCA GCCCGCCCGT CTTAATTGAC CGGTAAATTT CCAGTTCTTC TCCTCCTCGC 4440
ATCGGTTCCT GCATGGGTAG CTAGCTAGTA ACTCCGACGC TTCTGCTGGA TGCAAACACT 4500

FIGURE 5 (Con't.)

```
TGTGCATTTT CAGGCTGTGC AACCAGGAAT GCAAGTCGCC GATAAGGCAG ACCGCAGTGT  4560
CATTCGACGG AAGGCACGTA CGCACTACGA CTCTCACTAT CTGCTCATGC ATGCATTCAC  4620
CGCACGTACG TGTGATGTGC TCGCTCGCTT CCTCCTTTTG TGATGGTGTC TCTCTCACTT  4680
GCCCAGCACG ATCTTGGAGC CGCCGACGAC GGCGGATCTG GCGCGGTGGG ACGAAGTGGA  4740
CCCTGCTGCT TCCAGCTCCA AACCTGATCA AGCTGCTGCG CCCGCCGCCG GTGCGGGTGC  4800
CGACGCCGAC GCCGACGCCT GAGCGAGAGG ACCGTCGTCG CCCGCCGGTT CACATCGATC  4860
GTACTCCGTG CTGGCTGATT ACCTTTACCC ATTGGGATGT TTTGGTTCAG AGTCGCCAGA  4920
TCTAGTGTGT GGCTGAACGT TGAATGTTAG GATGCTGCTG CTTGTTATGC TCTGAGTCTT  4980
GAGTTCTCTT TGTTAATTTG CACCGTGGAT GAGATGAATA ACTTGACGTT GCAACTTTGC  5040
ATCCCATATA TGCCGTAAAT CTGCCGTCTG TTGTTTGTTC TGCGTTGTCT AGAATTAGTG  5100
GAGATGTGCT GGATACAATG TATGCTAGTC TATTAAACCG TGCTCCACTC TGAGATAATC  5160
GACCAACTTG TCTTATTATT GAAAGAACTG TGGAAAAAAC CAAAAAAAGT CGTTGTGGTT  5220
TTGTTTATTA TCAAATATAT TTTACATAAG ACTTAAAAGT TTTCATTTTT TCATGAATTT  5280
TTTGAATAAA CCGAGTAGTC AAAGCTAGGG TCAAAAAGGC AAACATATTA TATTTTAAAA  5340
TGGAGAGAGA GTACATTGTT TTAAGACGAA TTGTTTAATA CAACTCGAGA ATATTCTGAT  5400
ACATTAATCC TATGATATTA CCATAAAAAA CATTAATCCT ATGATAGAGT GTATAATTAC  5460
AAATGCACAA AGGTTCTTTT CATGTGAAAT CGTATTATAG ATAGGGGTCA TAGCGCGCCC  5520
TTGTCCCTAC AACTTACGAT GTTCATGAGT TAGGTTAGAA AAAGGTTAGA GCAAGTATAC  5580
TAAAGTGACA TATGCAGGCT ACAAGGAATG CCACATCAGA TTTTTGGTGA CGTTGAAGGA  5640
AGAAAAATAG AGGGAGAAAA AAGCGAACCA ATTGCGAAGG TGCCTTCTTC CAAGGGCACG  5700
GTCCATGGAG TGTGGTAGCC GACATCAAGG TAGAGGATTA TGGTAAAGTT ATTTGAGCAA  5760
GTGTCTGACA ACTAGCATGA AGGCTTAGGA TTTTCTAAAT GCATCTTTGA GCGCTATTGA  5820
TGTAGATGTT AATGATTTTT AGGGCTGATG ACCAAACCAA AGATGAACAT GGGAACGNAA  5880
GGAAGGTTAC TGAAAGTGTA TAGGCCCCTA GTTTAGTCTT CAGTGACTAA TGATAATATA  5940
TATTATTGTG ACTAACAAGT GTTTTATAGA AACANGGAAA GTTAGATCAC AATAATAGAT  6000
ATGATCAGGA TTATTATGTG GTACCCATCC CTTATTGATG AAAATCAATG GTTGGTTCTC  6060
ATAGGATAAT CGAAAAGGTT AAGGATCAAC TGTAAATGGA GTTGTTGGAC ACTTAGAGTA  6120
GTGATTTGAC CTTTTTTCTT TGGTAGTACT ATAAACGGAC ATGAAATGCG TAGCTTTACC  6180
TAAACAAGTC TAGTTAAGTA TGATGATGCA CACTTGTGAA TACTAGTGCT AGGTAAACCC  6240
ATGAGATCTC ATGTGAAGTT CGAAACAAAA CCTAATTCGA AAAGTGATTA AAACATGTGA  6300
CTTAACAATG TTGTAGTAGC ATTGGTCGAG TTTGATGGGC ACCTGATATG GGTCACTAGA  6360
CATGAGTGTG CCCTGTTGTG TTTGAGTGAA GCACTAGCAT ATCAGGTGTG CAACAGATAT  6420
GGTGCACCCA GGCAGGACAC CCAAAGAGCT TGCAAAATTA GCCTAAAACA CTTAGTGCTC  6480
ACCAGACATA TCTAGTGTAC TACTAGTTAT TCTCGTTATA TATGAACCCT ATTAGTTATT  6540
CTTGAATTGC TTCGATCTTT TACAAAGGAA GTAGTTTTTC CTTCATCTCC ATAAACTGTG  6600
GTTTTCCAAA GGCATTAATA ATAAGATTTA GTATATTAAA TTCAAAGTTG AGGTACTTTA  6660
TTATCGTGAA ACCAACATTA ATACTATAGA CTTAACTAAG GAGTCTATTG GTGCTTCCTT  6720
CTCATGTATT TTCTTCTTGA AGTGTTCCTT CATCTTGGTG CTAACGACGA CATTCAACAA  6780
TGTGTGCTCT TACTTGATTG GTTTGTATAT ATGGTGGTGT TCCTTTACTT AGTGGCAACA  6840
TACCTTATCG ATAACTAACC CTTAGTGAAA GAAATGAAAA TGTACATCCC ACTGGGAAAT  6900
CACTCATACC CCTAAGAGCT AACTTAATGG AACATCACTC ATAGCCCTAA GGGCTAGTTG  6960
GAAGTACTTT CTCATTTCCT GTATAAGGGC TAGTTCATGA TTCAACTTCT TCTCCATTTC  7020
TTGGTGAACT ATCTTAGCAC GATTCCTATA AAAACATATA CAACTAAACA AAGGGTGGTG  7080
GTACTGAACA CAGTGGACCC AAGCACTCGG AAATGGGAAG GACAAGTTGC ATGGAAAAAA  7140
CGACAGGCTG GGAACTATTG TGTCTTGTCA AGCGTGTTCG TCCAGCTATA GGACATGGGT  7200
ATTTATAGGG CAACTAGAGG TTGGTATCCT AAAATATGTC CAGACCCCTA GTTATCAACT  7260
ACGTTCCTAG ATAATACTGT ACAACAAGGT AATTATAGAA TAGTAAGTTT GTTATTCTAA  7320
CTCCACCCCG ACAGGTGGGT CCGTTGTCGC CCGGTTGAGA GTGGGCCCTG CTCGGCCAGG  7380
TCATTGGCAT TGTCCGTGCA GACGTGTTCC CAATATCGAG GCAATGAAGT TGTTTGACAC  7440
TTCTTCGGGA GTCGGCGTGA GGCCTTCGCT TGCTAGCGCG AACTTGCCCA CGAGCGTCCT  7500
CACCATGGGC CCCGCTGACA AGCTT
```

FIGURE 5 (Con't.)

>ZmFie2 5' upstream

```
   1 TTTTTCACAC CGTTACTGTC ATCTAACAGA AGCAGGTACA AACTTGTTTT TCGTTTTCAA
  61 GTCGAATTTT GAGGGGCAAA CCATAGTTGC ACTTCCATCG AGGGACAAAA ACACAATTGC
 121 CCCTTAACTT ATATAGTTAA ATATAGTTAA CGAGCTTGCT ACTGAGACTA ACAAGTCAAA
 181 ACTATTGGCT TGACCTTATA TTAGTTTTGT CTTACACTTT ACAATCGTTG ATGGCTGCTC
 241 TAGATCTTAT AAACTTAAGA ATATTATGAC TTTATCACTT TATTTGTAAT GGATGTATGG
 301 ATACTCATTG ATGCATTATT TATGGTATAA ACTATAGACC ATGAATGTAT GGTGTAATGC
 361 TATAGTATAT TGTTAGACTT GTGTACATAT ATATTATTTA TACTTAACTC ACAAACTTAA
 421 TGAGTCAGCT CGAACTTATA AACGACCTGA GTCGACCTGG CCTTATGGCT TGTTAAGATA
 481 ACAAGTCAAA CCAAGCCGAA CTGACTCGTT ATCCAAATCT ACACTTACAT AAACAAAACA
 541 TGATTTCAAA TTAAGATTGG TACAAAAGTG TTTTGTTTTA TTCAATTAAA CCCTACACTG
 601 TACTCTTTAT GTCAACAATA GTTGATGCTA CGACAAAGCA ATGAACATTT TATGGAGTAG
 661 TTAATTTTAT TGTCCTAATG TCAATTACTA TTGTTAGCCA AGGAATGGAG TAAGCCAATA
 721 AAGAGTACAT ATCTACGAGG AAATTTAGAT ATGTGCGTAA CTTTTTTAAT CGAGATACAA
 781 AATGTGCAAA ATAAGGGTCC ATGTAACATA CATATATTTC TTGTTTTTAT GGTAAAAGAG
 841 TGTATAAACT ATAAAGGTTG TTGCTTAGAA GCGGGATTTA ATAACATCGG TTTTATATTA
 901 ACCTTAAGTC CCTATGCAAT ACCTGTATTT TTTTCTAAGT ACATGGTACA AACACAAATA
 961 CACACATTTA AGCACACATA CTCACTTGCT ATGAGCACAC ACACGTAAAC CCTACTCCTA
1021 CTAGCACCTT CAAAAGACAA AATAGATAAA TCTTGTTGAC AAAGTCTATT GAAAAATATC
1081 AACGTCCGGT CTAAATCTTG ACAAAATATT AGCACTTGTG CCAAGTTAAG AAGTGAGCAC
1141 TTGAACGTAA GTGGTTAGAG GAACCTAACC AAGTTAGTTA TGTTCAATTT TTCATGCAAG
1201 TTAGCTTGCT AGTTTTTCTA TACACAAACA TTATATTAGC TTATACCATT GTTGGGAAAT
1261 TCTAACTTTA ATGATTTCTT TGAGAAATCC ATAAGAGCGA TAAAGAGGAG AGAGAGAGAG
1321 AGCAAGAGAT TTGTACATGT ATAAATACTA TCCATTTTCT ATTTAAGAAT CTAGACAAAC
1381 TAGCAAATAT AAATTTGAAA CATAATAAAG ATGGGCACCT GGCATCTCCT GGATATTAAA
1441 AGCGTACCAT TAAAGATATA CATAATTATT CACCTCTTCT AGGTATAAAT TACCCTACTA
1501 CCACATTCCC CTATCTCTAC AAACTCTCTC TCATTGACTC ATCAAGAGAG TGCCACCTCT
1561 ATCTCTCCTT CTCTCTTTTC AAATGTTCTA CAATTATCAA CCATCATACA ACATTCACCT
1621 TTCCTACCAA CCTTGTTGAT GCTTGTCTCA ACTTTCTCTT TACCTAGATC ACTCATATAT
1681 ATCCCTATTT CAAAGGCATT AATCATCAAA AACCTATAGA AAAATCCCAT TATCAACCAT
1741 GATGGAGTCT GATCGTGAGA AACAACAGTC TCATGGCAAG AAACAAGGTG ACCATGGTAG
1801 CAAGATGCAT GATTCTGATG GCAATAAAAA TGTGTCAGAT GAAAAGAGTC AAGAGTCTGG
1861 TGGTAAGGAA CACAAATCCA ATATAAAGAA ACATGAATCA CGTAGAAAGA GGTAAGACAT
1921 TCTCCTTGAA AATCTTGGCT TCAAACTCAA GTTAAATTTA TGTACACATG TTTATATAGA
1981 GTCTAGAGAT TTTGTGCTTA ATATATGCAT GCACATGAGT TCAAATAATT TCATAATAAA
2041 AATAAAAAAA TCAATATGAT CAGGAATTAA ACCATGAAAT TTTTAGAGAC ATCATCTAGA
2101 TTGAGTTCCA TGGTCATACC ATGATGGTTA TGTCATTTCT TTCCAATATA AAAAATTCCT
2161 TAACTTATAC TCAAAATGTT GATTGGATGG AACTTTTTCT ATAGAATTCC TTGCCACATG
2221 TTGTGTAACA ACCATTTGTA TTGGTTTGCG TCTAGTCCAC TTTTGTGTGT TGCTATTATG
2281 TAAATAATTA TTTTTCAAAT CCAAAGTTGT TCCTCCACAT ATCTAGAATA TATTCTAATT
2341 CTACAAGAAT TTAAAATGAA TTGTTAACTT AAGAATGCAT TGTTCAATAT ATTTATGCAT
2401 TTTCTCCCAT TATGATATAT ATATTCTCAA TATTTGGCAC ATAATAACTT GGAACATTCC
2461 TTACATTTGT TGGGTTGAGT GCTATATGTT TGGATTCATT AATTATTTAC ATTGATATTT
2521 TTGTAGATGT TTGTGTTTAC CCAATAAGAA AAGGCCATTA AGAAAATAAA ATGTTATTAG
2581 ATAGAGTTAG TCTTGACATG TTATATTCTT TTAATAATTG GATTTTGTGG TATTTCCAAC
2641 ACATTCCTTC CATTTAAACC TAACTCCATC TCTCTTATCT TCCTCTATCA TATACCTTAT
2701 CTTCTTTCTA CACTAACACT AATGCTTATG TCACTCCTAA CCTTGATGCA ACCTACCAAT
2761 AGTCAATTAC TGTTACGTTG CTAGAACCAA AGATTGGTCC ATTGGTGCAC AATCCATTAG
2821 TTCCTCCTTC TTGGGACTCT TCAACCATCC TAACTCCCCA AATGATTTCA AAAGTTTTCC
2881 CTACCATGTC ATCCTACTCC ATATCCAATG TCTACTGGTG CTAGATTCTA TCTACTGTTA
2941 GCACCAAACT AACCACAAAA TAATAATCCC TACAAATATA GGTGGAGGTG ATGTAAAATT
3001 AAGGGAGGGG CAATTGTAAA TGGTAGTACC ATAGATATCA AACCTTCTCA ACTTAGAGCT
3061 ATGTCTACAT AGTTCTAGTC CTATGAAGCA TCAACCATTT TCTTACTAAA CTAAATATTT
3121 TTAGAGGAAG GGGTGGATCC TTACTTTCAT CTCCATGAGC TTCCACCCCT TCCTATGAGC
3181 TTATCCATCG ACTGAAAGTT CCTCATTGCT GGAGCTTACC CGTTATTATC CCATGTCATC
3241 TGACTTTTGT ATGTACTATT ATCTTTGAAG TCGTAGGCAT GTGGTAAATT CCTACCTTAA
3301 GATCCATTAA TCCTCCAACA CACCCTTAAG ACCCAAACCA TAACGCCTAA ATCCAATTTC
3361 AACATATTTT AGGTGACATG GGTATATGTG ATATTAGTTA CTTAATATAG CAAGCTCTAT
3421 CAATGATTTT TAGTCAGAAA ATGGTTGATA TGTTTTTAGT GGTTGTACTA TAATTGAAGA
3481 GGCACATAGA GCAAGTTTTT AGACCATGAA TATATGGTGT AAACTATAGA CCATGAATGT
3541 ATGGTGTAAT GCTATAGTAT ATTAATTATT AGACTTATGG ACATATATAT TATTTATACT
3601 TAACTCACAA ACTTAATAAG TCAGCTCGAA CTTATAAACC ACCTGAGTCG AACTGGCCTT
3661 ATGGCTCGTT AAGCTAATAA GTCAAACCAA GTCGAGCTGA TTCATTATCC AAATCTACAC
3721 TTATGTAAAC AAAACATGAT TTCAAATTAA GATTGGTACA AAAGTGTTCT GTTTTATTCA
3781 ATTAAACGCT ACACTATACT CCTTATGTCA ACAATAGTTG ATGCTACGAC AAAGCAATGA
3841 ACATTTTATG GATTAGTTAA TTTTATTATC CTAATGACAA TTACTATTGT CAGCCAAGGA
3901 ATGGAGTAAG CCAATAAAGA GTACATATCT ATGAGGAAAT TTAGATATGC GTGCAACTTT
3961 ATTTTTTTAA TCGAGATACA GAATGTGCAA AATAAGGGTC CATGTAACAT ACATATATTT
4021 CTTGTTTTTA TGGTAAAGGA GTGTATAAAC TATAAAGGTT GTTGCTTAGA AGCGGGATTT
4081 TAATAACATC AATTTTATAT TAACCTTAAG CCCCTATCCA ATACATGTAT TTTATTTCTA
4141 AGTACCTGGT ACAAGCATAA ATACACACAT TTAAGCACAC ATACTCACTT GTTATGAGCA
```

Figure 6

```
    4201 CACACGTAAA CCCTACTCCT ACTAGCACCT TCAAAAGACA AAACAGATAG ATCTTGTTGA
    4261 CAAAGTCTAT TTATGGTATA AACTATATAC CATGAATGTA TGGTGTAATG CTATAGTATA
    4321 TTGTTAGACT TGTGTACATA TATATTATTT ATACTTAACT CACAAACTTA ATAAGTCAGC
    4381 TCGAACTTAT AAACGACCCG AGTCGAACTG GCCTTATGGC TCGTTAAGAT AACAAGTCAA
    4441 ACCAAGCCGA GCTGACTCAT TATCCAAATC TACACTTATA TAAACAAAAC ATGATTTCAA
    4501 ATTAAGATTG GTACAAAAGT GTTCTATTTT ATTCAATTAA ACCCTACACT ATACACCTTA
    4561 TGTCAACATT AGTTGATGCT ACGACAAAGC AATGAACATT TTATGGATTA GTTGATGCTA
    4621 CAACAAAGTA TATTGTTAGA CTTGCTAGAT TCTATCTACT GTTAGCACCA AACTAACCAC
    4681 AAAATAACAA TCCCTATAAC TATAGGTGGA GGTGATGTAA AATTAAGGGA GGGGCAATTG
    4741 TATATGGTAG TACCATAGAT ATCAAACCTT CTCAACTTAG AGCTATGTCT ACATAGTTCT
    4801 AGTCCTATGA AGCATCAACC ATTTTCTTAT ACTAAACTAA ATATTTTTAG AGGAAAGGGG
    4861 TGGATCCTTA CTTTCATCTC CATGAGCTTC CACCCCTTCC TATGAGCTTA TCCATCGGTT
    4921 GAAAGTTTCT CATTGCTAGA GCTTACTCGT TATTATCCCA TGCCATCTGA CTTTTGTATA
    4981 TGTACTATTA TCTTTGAAGT CGTAGGCATG TGTAAATTCC CACCTCAAGA GTCAAGATCC
    5041 ATTAATCCTC CAACACACCC TTAAGACCCA AACCATAACA CCTAAATCCA ATTTCAACAT
    5101 ATTTTAGGTG ACATGGGTAT ATGTGATATT AGTTACTTAA TCTAGCAAGC TCTATTAATG
    5161 ATTTTTAGTC AGAAAATGGT TAATATGTTT TTAGTGGTTG TACTATAATT GAAGAGGCAC
    5221 ATAGAGCAAG TTTTTAGTCG TTGTATTCTA AACAATGATT GATGTGTATA AATTTAATAA
    5281 ATTCATTGTT GCATCTTGTG TTTCATACAT TTGAAATGCT TTGTGCCTAA TCTATATGGA
    5341 TGAAGAAGTA AATCCTTCTA AACTTTTCCT TCCCTGCAAT CTTTTTAAAC ACACTCTAAA
    5401 CCCCAAATAT CTAATCCTAA CCTCTAAACC TGATTTAAAT TTTCTAATCT AGTCCATTTG
    5461 TAGTGCTTTT ATATTTAGTC CATTTGCCTT ATGTGCCTCT TGTGTATAAA TAGCGTAGAG
    5521 TTCTGTATAA TAGTCAACAA GTTTTGCCTT TTGTTGTCGG ATCCATTTTC AATCCTTTTG
    5581 TCTAGTTCAC CTATTGTTGT TGTGAAAAAA ATGTCACACA TTTTTTACTT CCCCCTATAC
    5641 CACATACTCC ATCACGGACT AATGATCTTC AAGGTATGTA TGCTCAGTTT AAATCCATGT
    5701 CTCCACATAC TCCATCTTAA GTTCAAGTCT CTACTTTAAG GTATGTAATT TTAAAACTTT
    5761 GACGTATTGT AATTCTATAA GGAGCAAATC TGAAAATTAA ATAAGGAAAA ACTGGTAAAG
    5821 GCATGTTTGG AAATCGGAAC GCAGACATTT TGTTGTTCCT ATGTTTTTCT TTAAATAAAC
    5881 TCATTCGTGT AAAATTTCTT CAAAATTCCT CTCCTTCGAA CAGATCCTTT TGCCCCCGGA
    5941 CCCCTTTCCT ACGCTTGCCC AAACCCACAA AACCCTCGCC GTCGCGCCGC GCGATTGCCT
    6001 CTCCGGCCGC CGCGAGCCCG CGACACTAGT AACGGTCTAC ACCACCAGAA TGACTGAAGA
    6061 ATTGAATTCC AGCAAATTCA AGCTTTTGTT TTAGCCAAGA TTTGAGATTC GATTTGAAGT
    6121 GTGGAAGTCC TTCCAATTTG CCAATCCTAT ATTTGATCTC TGCTGTGCTG CGTTAAATCC
    6181 CTAAACTTCA CAGCGCGGCG CCGGCCCAGC CACGCCGGAA GAGGTCGCCG CGTGAGGTCA
    6241 GTGTCCCCGT TGCTGCCGCC TCTAACCCGA AGCCTAGGCC GCTGCCGGTG CATAACAAGG
    6301 AGAATCAGGC GGAGGGGAAA GTAGCAGAGG AGGGGGCAGC AACTGAGGAG GGGGAGAAGT
    6361 ACCGGGCGGA ACCGGAAATC TTGCCGCTGC CGCCGGCCAT GGCGAAGCTG G
>ZmFie2  coding region
       1 AAGCTTTTGT TTTAGCCAAG ATTTGAGATT CGATTTGAAG TGTGGAAGTC

      51 CTTCCAATTT GCCAATCCTA TATTTGATCT CTGCTGTGCT GCGTTAAATC

     101 CCTAAACTTC ACAGCGCGGC GCCGGCCCAG CCACGCCGGA AGAGGTCGCC

     151 GCGTGAGGTC AGTGTCCCCG TTGCTGCCGC CTCTAACCCG AAGCCTAGGC

     201 CGCTGCCGGT GCATAACAAG GAGAATCAGG CGGAGGGGAA AGTAGCAGAG

     251 GAGGGGGCAG CAACTGAGGA GGGGGAGAAG TACCGGGCGG AACCGGAAAT

     301 CTTGCCGCTG CCGCCGGCCA TGGCGAAGCT GGGCCCGGGG CAGGGGCTCG

     351 GGTGCGAGGC GGCGGAGGGG TCGCTCGTGC CCAGCCGGAA GCGGGAGTAC

     401 CAAGCCCTGC GGCAAGCACA CTGAGGGGAA GCGCCCGCTA TATGCTATCG

     451 GGTTCAACTT CATGGACGCG CGCTACTACG ACGTCTTCGC CACCGTCGGC

     501 GGCAACCGCG TAAGCCATCG ACTGCTCTCT CCTGTCGTCC TTTTTTTGTT

     551 TCTACTGAGG TTTGGGGAGT TCTTGTTGAT TAATGGCAAG GTAAAACTAC

     601 GTTGTTTTTT TTTGTGATTT TGGTGGTCGG TTTTAGGAAG CGGTCGCTTT

     651 TGATTCAAAT TTGATCTAAA GCTGAGGCAT TCGGTTGTTT TTATTGGGGA

     701 CTTGAGGTCT GTAATGTTCC GACTATTGTG ATTTGTTTTG CCGAAACATG

     751 GAGTTTGCTA GTTCATTTGA TGAAAAGCTG CAACCTTTGA CAAAGAATTT

     801 GTATCACTTG GGAAAGTATA GTGAGGTGTG GGGAATCAGA TAGTACCAAT
```

Figure 6 (Con't.)

851 ATTACTTTGA CTATGATTAT AAGATAATCT TTTAATGTCC TTTGTAACGA

901 CCATGCTGCT TTTCGCTTAT CTTGCCTATT GATCTTGCAG GTGACAACTT

951 ACCGCTGCCT TGAGAATGGT AGTTTCGCTC TTCTACAAGC TTACGTTGAT

1001 GAGGATGTAA GAAAGACAAT GCTCAATGAC AATGCTTTTG CTTGCTGATT

1051 TAATATTGAT AATATTCTTT CTCTAATTCT TGTGACGCCT ATTTACCTCA

1101 GAAGGATGAG TCGTTCTATA CTCTAAGCTG GGCTCGTGAC CATGTTGATG

1151 GCTCACCACT GCTGGTGGCA GCAGGAAGCA ATGGGATCAT TCGGGTCATC

1201 AATTGTGCTA CAGAAAAGTT AGCTAAGGTA ATCTACCCTT ATATTTGTAT

1251 GTGTTCCTAT GGTAAACTTG AATGAAGCCT TATTTGCATA ATTCAATATT

1301 TCAGTTGTTT ATTTGACATA TATCACTTTA TTTATGATAT CTGATCCAGA

1351 AGGTCTTTTG GATTTGCTTT AGTTAAGGAA TGGTGCTTGC TACGCATTAA

1401 TACCATAAGC AAACTGTACC TTTTGCTCAC AGAATATTGT TAATTTTGAC

1451 TACTTCAGTA TGTCCGTTGT AGTAAAAACA AATCAACTTG GTGTATCTAT

1501 TTTTTCCTTG CTTATACATA GCCAGGAGAT TGGGCATGTG GCATGTCAAT

1551 AAATACTATC CTATACCATT TGATAGGACA CGCACTGTGT CTTATTTGGT

1601 AGCTCTGTTT ACGTGATTCT GCAGAGCTTT GTTGGCCATG GCGACTCAAT

1651 AAATGTGATA AGAACTCAAC CGTTGAAGCC TTCGCTCATC ATTTCTGCAA

1701 GCAAGGTTAT GCGATAGTCT GTTCTTAGGT TCATGTACCT TTTTATTTTT

1751 ATAATCTTTC TGAATTTTGA CACCATTTCA TATGGCATTA TCTAATAGGA

1801 TGAATCTGTT AGGCTATGGA ATGTCCATAC AGGGATCTGT ATCTTGATAT

1851 TTGCTGGAGC TGGAGGTCAT CGCAATGAAG TATTGAGTGT TGTAAGTAGT

1901 GCCTGCTATT ATGACATTGT GCCCTTCAAA AAAAACATTA TTATGACATT

1951 ATTTTTAGAA CATTACTAGG TTAAGGTGCC TTTAATATGG CGCACTCTTT

2001 CAGCTCCTGA TATTACCATT TGTTATTGAG CGTTACATCA GAGATAAAAT

2051 AAGGCTACCT AATGACTGCT ACTGCTTTTG TACTTTGATT ACATTAGTCA

2101 TAAATGTACT GATGAATACA TTATTTTGTC TTAAGGACTT CCATCCTAGT

2151 GATATTGAAC GTTTTGCAAG TTGTGGCATG GACAACACTG TGAAAATCTG

2201 GTCAATGAAA GGTTAGAAAG CTACTTCAAA GTTGCTTCAT ATTTGCATGT

2251 TGCGTGTCAT TGAGTTCACC AATGTTGTCG CAGAATTTTG GCTATATGTT

2301 GACAAATCAT ATTCATGGAC TGACCTTCAT CAAAGTTCCA CAAAATATGG

2351 CCAGTTTCCA GTATGTTTCA CAATGCCTAT ATCCAATTAT CCTGGCAAGG

2401 TCCTGTTGGT GTCTAATCCT CATGCCATCA GACTGACCTG TTTCTTTTTG

2451 TTTCAGGTCT TGATTGCTGC AGTACACTCT AACTATGTTG ATTGAACAAG

2501 ATGGCTTGGT GACTTCATCC TATCAAAGGT GAAATTTCTG ATTCGTTTAA

2551 ATGGATACAA ATTTCTGTAG CACGGTTGTC ACTCTTTTGT GGGTTTGACA

Figure 6 (Con't.)

2601 TGCCACTGTC TTGGTTCATC TATTGCTGTA CCGTGCAAGT GTTCAGTTTT

2651 TTCAATCTTT TTTCTCAGTG CTTAATGAGG GGAGATTCTA TTTGCAGAGT

2701 GTTGTCAATG AAATTGTGCT TTGGGAACCG AAGACAAAAG AACAGAGTCC

2751 TGGGGAGGTA ATTCAGTTTA ACTTTCCCAG AATTGTATTC CTATTATAAT

2801 GCCATATATT TACGCACAGT TGTAAACTAT TTCCAGATCC TTAGATTTCA

2851 AGGTACTGGC TGCCAATATT AAATATGTTC CACTGAAGTA ATATGATTTT

2901 CTGTTGCCTC ATAGGGAAGC ATCGATATCC TTCAGAAGTA TCCTGTCCCA

2951 GAATGTGACA TTTGGTTTAT CAAATTTTCA TGTGATTTTC ACTTCAATCA

3001 GTTGGCGATA GGTAATATCT CTCATCAGGA TTGTTTCTGG TAGAAGTTTT

3051 ATTTAAGATT TTTTTTGCTC TGTAAAATTT CACACACGCA CACATGCACC

3101 CCCACACACA CACACATGCA CGCACACCCC CACCCACCTG CACGCGCGCG

3151 TACACACACA CCGCACACAT ATATATGACT TTTTTTCCCA CACAAATATT

3201 TGCTGTGTGA GATATCAGCA AATAAATTCG TATGTTTGAT TATATTCAGA

3251 GATATAGGAA AATTGAGTGC TCTAATACCC CATCCACTAC TTCAAACAGG

3301 CAACCGTGAA GGCAAAATCT ACGTGTGGAA AAATACAGTC CAGCCCTCCT

3351 GTCCTCATTG CTCGGTAGTT TTCACTGGAA GAGTTTCAGT TATTCTTGTC

3401 TCCCACTTGT ATCGTCGCAT GCTTCTGGAT GCCAATGCTT CATCATTTTC

3451 AGGCTGTATA ATCAGCAGTG TAAATCGCCG ATAAGACAAA CTGCAGTGTC

3501 CTTCGATGGA AGGTACCTCA CTCTAATCCA TGCTCAATTT GGTGTACTGT

3551 CTATTCTAGC ACTTGCTTTT TTCTTGGTTC TGCTTGAGAA ATTCTCGATT

3601 GCATGTCATA TGCTGGTGCA TTTTCTTTTT TCTGTTTCCG TGGCGGATTG

3651 GTAAAATGCG ACGATGCCTT CCTTATCTAG CACAATCCTT GGAGCTGGTG

3701 AAGACGGCAC CATCTGGCGG TGGGATGAAG TGGACCATCC GAGCTCCAGA

3751 AACTGAAGAA GTGTTGCCGC TCAATGCTGG ACTGATGGTT ACGCTCGGTT

3801 GGGGTTGTGA TGGTTGAATC CGTTGGCGGA AAGTGCCACC TGGTGTTTTT

3851 TTCTAGTCAA AATGGTTGAT GTTAACAGAA TATTGAATGC TTCGAATGTT

3901 GAAAGTTGGG ATGCTTGTGC TGGTACTCTG CTCCGCGGAC GAGTGAACTT

3951 AGTTTGTTGC AACTTTGGGA ACCGTTGTCA TCTGTTTGTT CTGCATTTCT

4001 AAAAAGAGAG CAAATTTCAG GATACATGTT CTTTTTTTTC AGTACAGGAA

4051 AACTAAGGTT GAGGTATTGC TTTGCAATTT ACTCTCTCTC TCTCTCTCTC

4101 TTAAAAAAAC TGGATCTTGC TTCAACGATG CATTCCTTGG GTCATCGGTT

4151 TTACTTTTGA AATCTTGATA GCTGGGCCTA AAGTTACCAA GCCCACTAGT

4201 ATCAGAAGTA ATAATATGAT GGCTCCTCCC CTGCCTTACT GTCACGTGTA

4251 AACTTTCGAA ACTAGCAGGA CTGTAGCATT TAGCGAGCTG GTTGTTTGGG

4301 TTAGAGCTCA GCGTCGCAAC TTATGGTACC GAGGTCAGTG TCAAGATCTA

4351 TGGCACCATG GTTCAATCAC AGTTTTAGTC CCACCAAAAA TATAAAGGTG

Figure 6 (Con't.)

4401 AAGTTTCGAC AAAAAATGGC TAGAATAAAA AAAAACAGGT CCACATACTG

4451 AGGAGAACAC ATGACAGATT CACCAAGGAT TTTGAATTGA AAGAGGCTAA

4501 TGATTGACAG GATTTGATCT TCAATTCCAC CTCCCGTTGT CCTGCTTCTA

4551 CTCTAAAGTT CAAGCGTGGC TCAGTTTGGC TATCTGTTAT AATTTCAAGA

4601 AATCCTGATT TCTGTTAGCA GTTTACTAGG CTATTAGGAG GAGCTGGGAC

4651 AAAAGAAAAA CGAGAATTGA CGAGGACAAA TTCGCAATTA GTTGGGAAAT

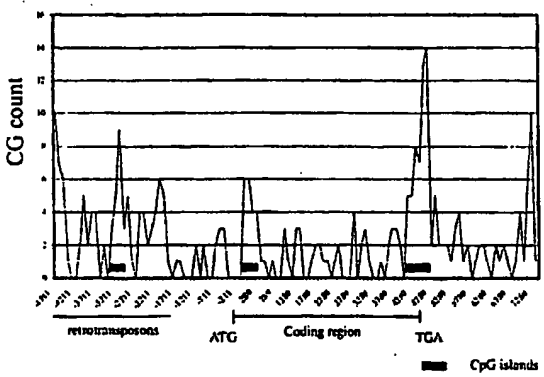
4701 TGGGGGCACA ATTTTCAATG CCCACAAAAT TCACTCCCCC TACNTNTGCG

4751 GNGGAATGGG GTCANNCCTC ANTGTCCCCT GTTNCCGGGA CAAGTNTAAC

4801 TAACACATTT CCNNATTNNT N

Figure 6 (Con't.)

CG distribution of *ZmFie1*



CG distribution of *ZmFie2*

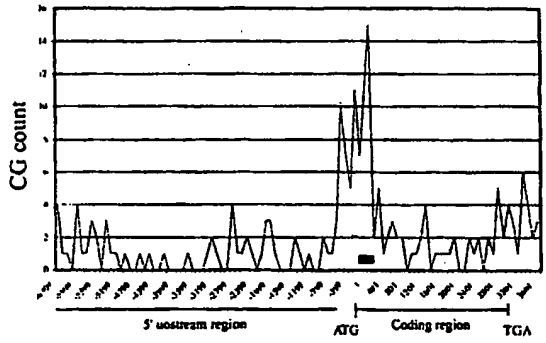


FIGURE 7

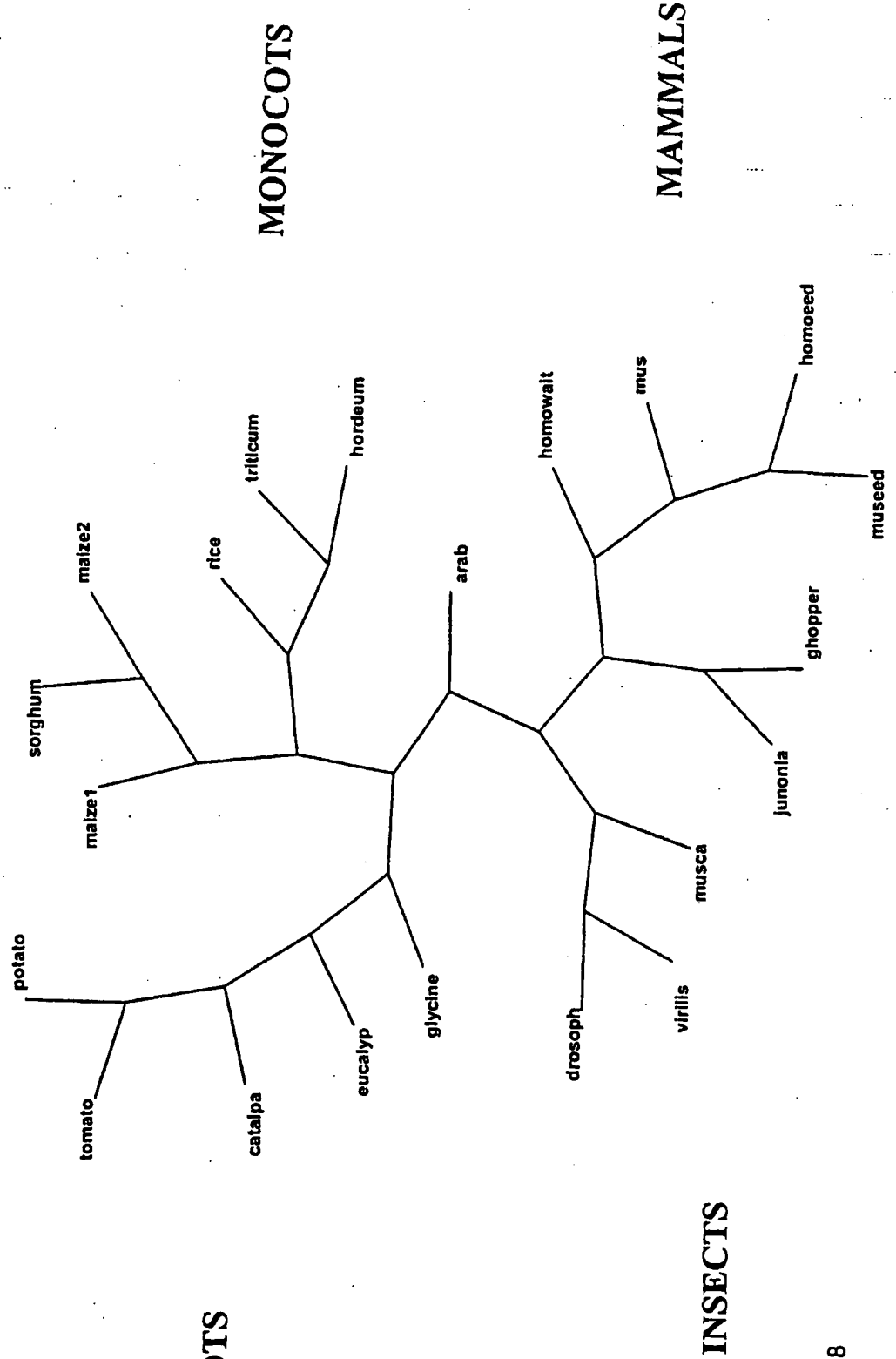

MONOCOTS
MAMMALS
DICOTS
INSECTS
homoeed
mus
homowalt
triticum
hordeum
museed
maize2
rice
arab
ghopper
sorghum
junonia
maize1
musca
potato
virilis
drosoph
glycine
eucalyp
catalpa
tomato

FIGURE 8

HpaII Restriction maps of ZmFie1 and ZmFie2 genes

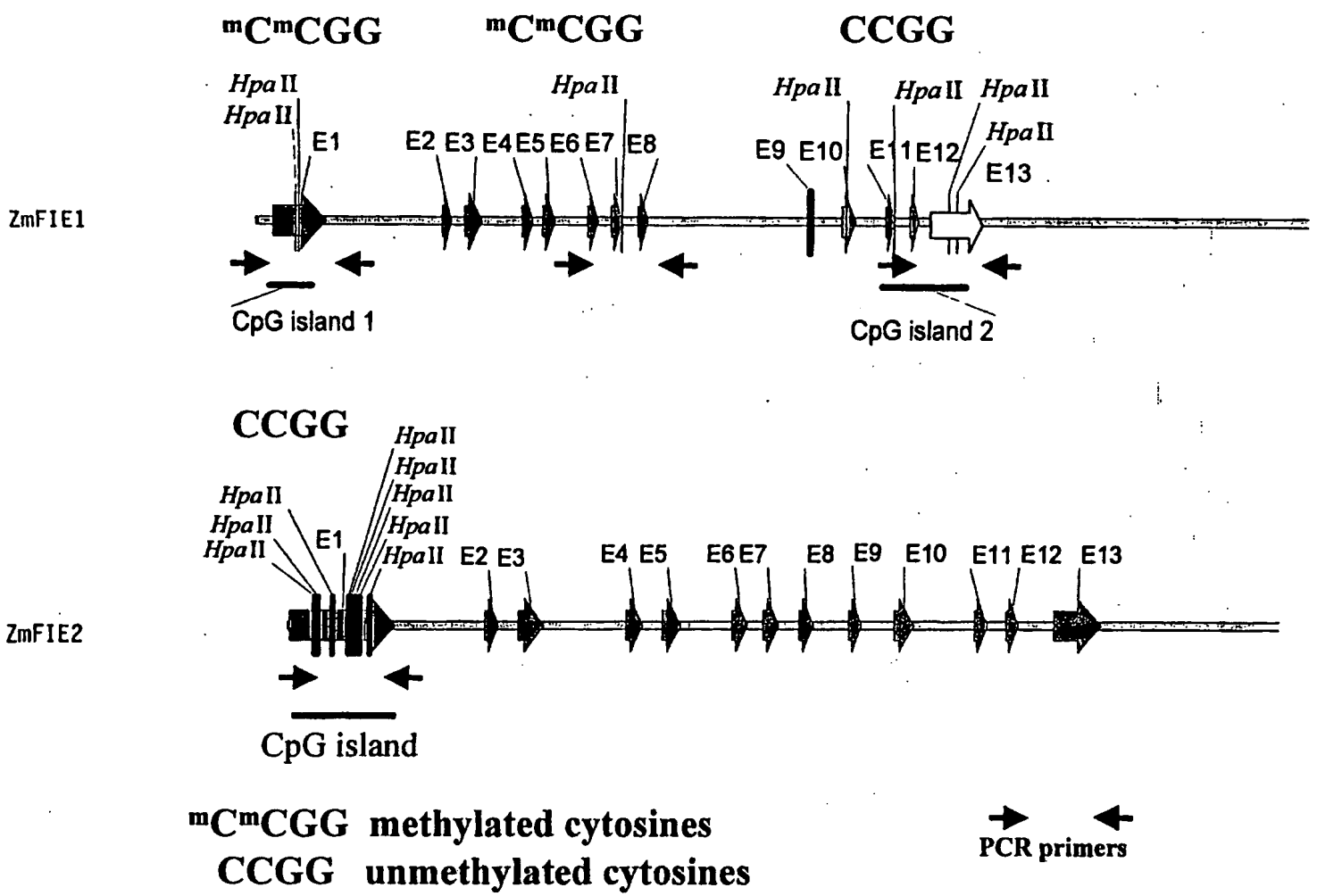


FIGURE 9

Primers for PCR amplification of ZmFie genes across CCGG sites

Gene	Biocode		Primer sequence
Fie1	66766	Fwd	AATTAACCCTCACTAAAGGGCGCCGCCACCATATAGAACCAC
Fie1	66765	Rev	GTAATACGACTCACTATAGGGCATTGCAACTGGCGATGGC
Fie1	35573	Fwd	ATGAGATAAGGACTCATGCCTCGAAGCCA
Fie1	44446	Rev	CCCACTCACGTGCAGGTAGAAAG
Fie1	53924	Fwd	AATGCAAGTCGCCGATAAGGCAGACCGCAG
Fie1	53926	Rev	CAACCAGCACGGAGTACGATCGATGTGAA
Fie1	53925	Fwd	AGGCGAGATCTATGTCTGGGAAGTGCAGTC
Fie1	34971	Rev	ATCGGCGACTTGCATTCC
Fie2	38860	Fwd	CGCGACACTAGTAACGGTCTACACCA
Fie2	36675	Rev	CGCGTCCATGAAGTTGAACCCGATAG

Fwd – forward primer; Rev – reverse primer

FIGURE 10

Single nucleotide polymorphisms for use in determination of DNA methylation of *Fie1* alleles

FIGURE 11

SEQUENCE LISTING

<110> Pioneer Hi-Bred International, Inc.

<120> Imprinting in Plants to Control Gene Expression

<130> 1487-PCT

<150> US 60/363,861
<151> 2002-03-13

<160> 6

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 17
<212> DNA
<213> Zea mays

<400> 1
gatctagtgt gtggctg 17

<210> 2
<211> 24
<212> DNA
<213> Zea mays

<400> 2
cgtgaaggca aaatctacgt gtgg 24

<210> 3
<211> 29
<212> DNA
<213> Zea mays

<400> 3
cattacgtta caaatatgtg aaccaaacg 29

<210> 4
<211> 31
<212> DNA
<213> Zea mays

<400> 4
cagaacaaac agatgacaac ggttcccaaa g 31

<210> 5
<211> 13031
<212> DNA
<213> Zea mays

<220>
<221> unsure
<222> (11384)...(11481)
<223> N = A, T, C or G

<400> 5
ccgatcattc gtttgttcga tcatttgatc gttcatcgtt cgttcatagt tcctattcat 60
cgttcatcgt ttgttcatag tacttattca tcgttcatcg ttcgttcata gttcctattc 120
atcgttcatc gttactattc atcgacacta ttcaccatcg ttactattca ttgttactat 180
ttaccggctc tattcgtcat cgttactatt catcgttgct atttatggta gctttttcgt 240
tgttactatt catcgatcat ccgatcgccc caaatttcaa ctactcatcc atcatgttgt 300
ccagtccacc taagaccagc cagacccata ttccagtcat acgaactcct gtgattgtga 360
ttttccttcc agtagggaac ctcccatctg gtcacccatc ctaggtttct ccaagttgag 420
catgcttaac tttgagattc ctttgaacca ggcttccaaa ctcagattcc aataattctt 480
gtttctaaat tcttatcaaa ctattcccta tccaaccatg tcatccctta agcctggtcc 540
atattccaga aaactcccaa aatactcttg tcccatattc tgcatataac tctcctgttc 600
atactaagtc agacgattca ttcgtcacta ttctcaccaa cagtgaactt cactgtgcta 660
caccacatac actcagctat aaatacaccc agctaccctc tccctctcca cacacactca 720
acaccctcag ccaaggcaaa cacctcaccc actcagttac tccgctctac cggctacacg 780
catagtgtcg cttcgcctcc agtccaccct cctggtaagc acctccgctc caccaccagt 840
aatatcacaa caccacatga cacagattct actcaagact ctacccatcc atatatcgct 900
attctgacca ctatactaaa tatttgttgg tatacttgct ggtttgtatg tttgcttgtt 960
catgttgcat agttatcgga gcgttcgtgc catcacgtgg aggccagatc tgcaagtcta 1020
cgccaggcgg tggagccaga agccagttcc gcgagctctc cttccccctt cactggataa 1080
gcacagcaag ctcactggat ccctttgatg cataaattac ctatgatttt tcaaccacaa 1140
ccctcagcct gttattttat gcataatatg attttgagac aagttattat ggccacccag 1200
ccgcttgtcg caatcaatcc ttgatatatt tgttacaaat gatttgagaa aaggtgtgag 1260
ttttcaaaag aaaatgcttt tcaaaatgtg tatgatgaag ggttttcacc cttatcacct 1320
tttaataggg atgatcaagg actccctggt ttaggggagg gcctaaggtg atggctcagc 1380
tggtttaggt gtgagcagaa ggattgtccc ctcacataag gaccgatttg tcatccgtca 1440
ctacctgtac tcatgataag tacaaccact cgagactgta tgggcaatca ctcaatctga 1500
actcgtacgg tccaacccta gggttatgaa ggctggggag caccgggagg ataaggaggg 1560
agaatgtttt gtccggtttg gacatggcgg tggcctgact ccttccggta taaccgttaa 1620
ggtaaggacg tgcgaggaaa gaaagagatc cggcattcgg gcctcacgac ggtgagatcg 1680
cagaaaccag actagtgggt aaagtgtacc cctctgcgca gagtttgaaa acctattcga 1740
atagtctgtg tccacaggaa tggacgagtc tggtgtggta tgacaattag tgttttgttt 1800
tcaaaaaaga atgtgcgttt gagaaaagtg gtttttaaaa ggtccggcgg ttgagccgtg 1860
agctatggtg gacgggaagt ccagtagctg tttttgaaaa cgaaaaccag tgggaaactg 1920
ctgagatacc tggatggttt agtccagggg attttgttct aatattgaaa aaaaattctt 1980
gctcctttgg gagaggatgc gctttgcaaa atacaaaatg ttttacaaaa taaccctgca 2040
taaaatattg ttgtttctgc aaaatatcct gagctccaca tattccatgc attatatctg 2100
atttccccat tccgcgggtg atggtgggct gctgagtacg tttgtactca cccttgctta 2160
tttgttgttt ttcaaaaaaa ggagatcggg taagagttac gactgttccc aaccttgcct 2220
gtggttgttg gaccgctgat ttgcttcgct gcgtatatcg ggctgcttca tccccactct 2280
gatgatatgt cccaagttgt ggaccaactc ttaaagttga tcgccacctt tataggtttg 2340
tctcgtttaa gcagatctgg aatcatttga tgtataaatg tgtttactag cctcctggga 2400
ctagtaattg tatcacattt gagtcctaga ggatcgggac gcttcaatga tcaatgggtg 2460
gatcacaata gtcggttata atggctatat caacagttat aatcacatta aatgtgtcat 2520
cagatgttag ataaagtctg tcgtggatga tctgtttgtg cttctcgacg gtccatgagt 2580
gacgctaaaa ttcatttac caaacctagc accttcgagt tggtctgatc ttgaatagtc 2640
agacggttca cgactgaggt tgaacgatcc acgcaaggtg ttggacgata ctttcttttt 2700
ctttggatgc tccgtagtag atgtgtcggt tttgacatag ttcctgtccg aactccatac 2760
agtccatagt agatgtgtcg gttttggtac tctagacggc ccgagtcagg ggtctggaca 2820
gtcctggact tgctgagttg aggtttgatc tttctttagt tatttcttac atacctatgt 2880
tcatacactt agcaaactag ttagcttcac caaaacaagt gtggaaaaag gtttttaggc 2940
caatttccct ttcaccttta taactaccta gttacaaagt agagtttgat agtccctaag 3000
tatgtcaatt cacatcttga gtacatgcga caatctcatg tctaaggata catggtacag 3060
gttgcaagaa gaaaattgtc acaatatctc atgttgggtc agtacagact catgtcatac 3120
atgcacccat attattagtt ttacatctcc atgtccatga cttacgaaac atagtcatca 3180
actaatacat atgatagtca ttgactctaa ctagggacat cttctagaac aaccatacaa 3240
gaaaagagtc tcacaaacaa ttcacataat tgctaatcaa tacaaggtgt ccttcacaga 3300
tattcaatta aacaatatat catggatgca acawaatatg ctcatctcta tgattatctc 3360
tagggcatat ttctaacaca atgacatgtc taagtgtagt atgtcaaaac atggatagta 3420
atatagatgg taagaggtca tttttattaa tataattaac aaagatagat agggtgacca 3480
attttgtaaa agcaccattc atagactttt agtgggaggt ggatgctcta cccgcctccg 3540
taaagccaaa gtggttgcat gcaaattgyt aggatatagt aatgcaagga accaagctaa 3600
ggcatgtaag tgaaacccaa acaagaagtt aagaagcttc caaaatgaac aaagtacaag 3660

aatgaagcta aaagagaaac tttcagcctt ctccaatctc cagcaagatc ccttcgatag 3720
atggtatcta atttttcct actatgaaaa cctatatcac ctagtagaat agaggacaaa 3780
gcttacgcct actatatata tccaatatgt atagttagat actaagttct tttttctctt 3840
ctcttcattc acttttcaac taggtttgga attaagtttt tggattggca tagacaatgg 3900
catggttgta taggtgttct taaccatcac agttatgagt ttgacttgtt ttttatattc 3960
aagttacaag gtcatttgt gctagccaca gcctagcaat cgaggggcta cacatgtgga 4020
ttaaggacaa ggcccaaccc atgtacgatc caaggacacc cttgtaattt ttatactcat 4080
caaggattag ggggaaataa ctcccttcta tataaaggtc tttccacttt gcttctcact 4140
ctcccttatt aggttaaaca caaaatgtgc atcgccgccg ccaccatata gaaccactta 4200
tcacgaaccg ccgccatcac atccactgcc tcaactagtg ttaccaccta tggttcattg 4260
ttgtgtctgc ttcttgtagc actgttggtc tacaaacatt catatttctc tcaacatctg 4320
gcacaggtaa gcccataagc cctaacccta gatctccata tttagttatt tcagttcttg 4380
atgagcaaat atgaaactaa attagtttgc taataagaaa tttaactact tttcctcttg 4440
aagacctcct atccctatat gaacccacat ccaaaacccc tctagcaaag tgtggctagc 4500
tttcccatgc catgaacctt caacaatgat agtatcagta atgcacttcc ataaaagggt 4560
tcatatttaa ttttagtttt tctttttggt gttttaatta agctttgaga cttgatttga 4620
agtattaaat aaacccttca aatttctttc taactttgat aatacactat tcaatgacaa 4680
tgcacttcct taaatcccta tacttcacag catgccgcct tccaaagcac gccgaaagag 4740
gtcacttcgt gatatcactg ccaccgttgc cactgggcct gttgccaact cgaaacctgg 4800
ctcatcatcg acgaacgagg ggaagcaaca tgacaagaaa aaggagggtc cacaggaacc 4860
ggacatccca ccattaccgc cggtggtggt gaatatagtc ccacgacaag gattaggatg 4920
tgaagtagtg gaagggctac tcgtgcctag tcggaagcga gagtacaagc ccaatagcaa 4980
gtatactgtg ggaaatcacc cgatctatgc catcgggttc aatttcattg acatgcgcta 5040
ctatgatgtc tttgccatcg ccagttgcaa tagtgtaagc aaccgacttc tccctacctc 5100
ttgtttgcta tccttttatc ctattgaggt ttggggagtt ctatatggtg aacgaaaatg 5160
gaagttatga ttttggtggg attggatctt ggtttataac tagaaaagga tttgagtaca 5220
ggttatgatg tgtggcttta tggtagggaa acttaatatc ttttcctatt ttgtttttg 5280
gcatcacgag taatggtttg ggaaataaaa gggaaaatga tttaaaatta tttctcaata 5340
gagcatgccc ttttacatag ggacatttta gtcattttac acacacttta gtcattttac 5400
acaccgtaat tatgtcacaa tcaaagaatc attccttggt tcaattgaat gagatgattc 5460
aactagttca catctctata cctaacaata tagttttca taactaaagc tttgagactt 5520
gatttgaagt attaaataaa cccttcaaat ttctttctaa ctttgataat acactattca 5580
atgacaatgc acttccttaa atccctatac ttcacagcat gccgccttcc aaagcacgcc 5640
gaaagaggtc acttcgtgat atcactgcca ccgttgccac tgggcctgtt gccaactcga 5700
aacctggctc atcatcgacg aacgagggga agcaacatga caagaaaaag gagggtccac 5760
aggaaccgga catcccacca ttaccgccgg tggtggtgaa tatagtccca cgacaaggat 5820
taggatgtga agtagtggaa gggctactcg tgcctagtcg gaagcgagag tacaagccca 5880
atagcaagta tactgtggga aatcacccga tctatgccat cgggttcaat ttcattgaca 5940
tgcgctacta tgatgtcttt gccatcgcca gttgcaatag tgtaagcaac cgacttctcc 6000
ctacctcttg tttgctatcc atttatccta ttgaggtttg gggagttcta tatggtgaac 6060
gaaaatggaa gttatgattt tggtgggatt ggatcttggt ttataactag aaaaggattt 6120
gagtacaggt tatgatgtgt ggctttatgg tagggaaact taatatcttt tcctattttg 6180
ttttttggca tcacgagtaa tggtttggga aataaaaggg aaaatgattt aaaattattt 6240
ctcaatagag catgcccttt tacataggga cattttagtc attttacaca cactttagtc 6300
attttacaca ccgtaattat gtcacaatca aagaatcatt ccttggttca attgaatgag 6360
atgattcaac tagttcacat ctctatacct aacaatatag tttttcataa ctagaattct 6420
taaaaagaat taatatgaac ctaaatatta tttcactttc ttgcccctta taatataata 6480
catttgtcac tcccattttg gcaagggtgg tgggtatttt gggggatgga atgttactat 6540
ttttaatttg attagaagct ataagctttg gctatatttt tattaggaat ttgatgttca 6600
ttttcaatat attgtgatct attttcttaa aatgtgaatt tgttgtgtat tttgattagt 6660
tcgatgaaga gtgtttataa gatatgattt ttaaattctc ttacgacgaa acaatattat 6720
gttactttca tctattcatc ttgaggaatc acctacctca cttcttgatc ttgcaggtga 6780
taatttaccg atgccttgag aatggtggtt ttggtcttct acaaaattat gttgatgagg 6840
atgtgagaaa gacaatgcct ggtgcatgtg gttgttaatg ttaatttgat aatatgcttt 6900
tatctaatgt ctgtggtgcc tatttatctc agaaggatga gtcattctac actctaagct 6960
ggaccatcga tcaagttgat agctcaccgc tgttggtggc cgctggaagc aatcggatca 7020
ttcgggtcat caattgtgct accgaaaagt tagataaggt ccctgccccc gtgcttactc 7080
tatgtttgta tggaaaagtt gattgaacgt tgatgttcac atatcaatat ttcagtagtt 7140
tagttgaaat acaatttatt tatgctctct attcttgaac atcagttgac tttgctttga 7200
ttaagcaatg gtcttgctca tacaatattc taggagttga atattcaata tgcctgttac 7260
atgatagcaa atacatagtg aactaggaca tgtactaaat atttaatttc cctttatgac 7320
attctctaga gcttagttgg ccatggtggt tcaatacatg agataaggac tcatgcctcg 7380

aagccatcac tcatcatttc tgccagcaag gttagtaata aatttgtcgt gtgtcgattt 7440
ttttacactt tttaacatga cattattcta taggatgaat ctattaggct atggaatgtc 7500
catactggga tttgcatctt agtctttgca ggggctggag gccatcgaca tgatgtgttg 7560
agtgttgtaa gtatcgattg catcttgtct agacattgtt ttaaatatca cttgccccga 7620
agataacact cattagaatt ctaatgttac catttgttat tgagcatgcc aaatttcaat 7680
tttaacatca tagataaaat aagaccccac aattactttt actgtttatc tacttccatt 7740
acattaggca taaagttact gataaaaaag acaatctttt atctgaagga cttccaccct 7800
accgaggttg ggatttttgc aagttgtggc atggacaata ctgtgaagat ttggtcaatg 7860
aaaggtttgg gaactacttt aaactagctt catgtttaca ttttgtgttg tatgttgcat 7920
atcatcgaca aatattgcca atgttgtcac agaattttgg atatatgttg aaaaatcata 7980
ttcatggact ggccatccat caaagtttcc aacgaggaat atccagtttc cggtatgtta 8040
agtagctata atcaccttgag ctcctttctt tttttgcaaa ctattgttgg tgttcagttt 8100
tcatgccatt caagcataca tgtttctttt cttttaggtc ttgactgctg cagtacactc 8160
tgactatgtt gattgtacca agatggcttg gtgacttcat cctatcaaaa ggtaaattct 8220
tcatttgtta aatggctata catttttta taaaggaaat tttttattaa tttcaagcac 8280
tttagattga aataatacaa aatcttaaaa aacatttttg gcctccattt aaacaagcac 8340
aaatccaaca aaaatgagta aaccaaccca ttctagtgaa tattaatgca taaactagat 8400
tgctacccat atgtctagaa aaagtagcct tgaccgcgta tcttaattgt caccatgccg 8460
ccacaaccaa accgtgcaaa tatggttttt ggagaatgga ccaagtaaga aaccaatcaa 8520
taattgagta tatagcatgc acaggagaaa tagatctctt attttcaaga acaatggtat 8580
tttttattaa ccataggacc aacaagtagc gactacccat agcaaaacta atggcttcag 8640
attattactg gttgttgaag tgtatacgtg gtttgcctac tttctcccaa tagtttaagc 8700
ttttggattg aatcgattag tgcgttcact cttacatggt atcaaagtta gcaattttgg 8760
gtttgaatcc taacggaagc tttatttgtg acttcacctc ttgttttcca tttccttttct 8820
acctgcacgt gagtgggggt gttgaagtgt ataagtggat tgcctacctt atcaaccttt 8880
tggattaaac tggttattgg ttagtgtgtt cactcctaca cctaagtatg aggtttagtt 8940
atccagtagc caattagatt atgcacagtg gacacttcac atgtgcaact agcactcaaa 9000
acataagtct ttaattgtct catcttatga caaaacaaca tatttcacta ccattctata 9060
acatcttgat ttgtacatca gtcttgttaa tgctaaatag tgagatttga tcgtcaattg 9120
gccagttgga tgtaaattcc agtgaaatac atcttgacct tgggttaaat ggacattagc 9180
aatgtgtggg aacaaattgt tggtttgggt acaccaaact gttggttttt aattagtaga 9240
ttagtttgta acacatttcc ttttatcagt gttagtattg gtttattatg catagggaag 9300
gatctgatat gtgataatta acatggattt gcagagtgta aagaatgcag ttttgctttg 9360
ggaaccaaaa ccagacaagc gtaggcctgg ggaggtgaca cgctttacct tctcgtcccg 9420
aattctgcac ctatttttat attactatca tactcatcta cagtttaaaa cttgtcccgc 9480
aatcttttca gtttctgagc actaaattta tacctctgaa tcagtatagt cgttttctct 9540
ttgttcgtat aggggagtgt tgatgttctt cagaagtacc cggtgccaaa gtgttcatta 9600
tggtttatga aattttcatg tgatttttac tccaaccaga tggcaatagg taatgccttt 9660
aattttgtga agactgtttt ggcactaaag ctttacgtac gtaatattag ttttatatct 9720
tgtacattga tggaaaatag attgctcaat atctatatat atgactatat cttgggttag 9780
attctaagga acaaactctc ccagagtacg gttctgaata acaaccatct gctgctgctg 9840
cttaatgcga acaggcaaca ataaaggcga gatctatgtc tgggaagtgc agtccagccc 9900
gcccgtctta attgaccggt aaatttccag ttcttctcct cctcgcatcg gttcctgcat 9960
gggtagctag ctagtaactc cgacgcttct gctggatgca aacacttgtg cattttcagg 10020
ctgtgcaacc aggaatgcaa gtcgccgata aggcagaccg cagtgtcatt cgacggaagg 10080
cacgtacgca ctacgactct cactatctgc tcatgcatgc attcaccgca cgtacgtgtg 10140
atgtgctcgc tcgcttcctc cttttgtgat ggtgtctctc tcacttgccc agcacgatct 10200
tggagccgcc gacgacggcg gatctggcgc ggtgggacga agtggaccct gctgcttcca 10260
gctccaaacc tgatcaagct gctgcgcccg ccgccggtgc gggtgccgac gccgacgccg 10320
acgcctgagc gagaggaccg tcgtcgcccg ccggttcaca tcgatcgtac tccgtgctgg 10380
ctgattacct ttacccattg ggatgttttg gttcagagtc gccagatcta gtgtgtggct 10440
gaacgttgaa tgttaggatg ctgctgcttg ttatgctctg agtcttgagt tctctttgtt 10500
aatttgcacc gtggatgaga tgaataactt gacgttgcaa ctttgcatcc catatatgcc 10560
gtaaatctgc cgtctgttgt ttgttctgcg ttgtctagaa ttagtggaga tgtgctggat 10620
acaatgtatg ctagtctatt aaaccgtgct ccactctgag ataatcgacc aacttgtctt 10680
attattgaaa gaactgtgga aaaaaccaaa aaaagtcgtt gtggttttgt ttattatcaa 10740
atatatttta cataagactt aaaagttttc atttttttcat gaatttttg aataaaccga 10800
gtagtcaaag ctagggtcaa aaaggcaaac atattatatt ttaaaatgga gagagagtac 10860
attgttttaa gacgaattgt ttaatacaac tcgagaatat tctgatacat taatcctatg 10920
atattaccat aaaaaacatt aatcctatga tagagtgtat aattacaaat gcacaaaggt 10980
tcttttcatg tgaaatcgta ttatagatag gggtcatagc gcgcccttgt ccctacaact 11040
tacgatgttc atgagttagg ttagaaaaag gttagagcaa gtatactaaa gtgacatatg 11100

caggctacaa ggaatgccac atcagatttt tggtgacgtt gaaggaagaa aaatagaggg 11160
agaaaaaagc gaaccaattg cgaaggtgcc ttcttccaag ggcacggtcc atggagtgtg 11220
gtagccgaca tcaaggtaga ggattatggt aaagttattt gagcaagtgt ctgacaacta 11280
gcatgaaggc ttaggatttt ctaaatgcat ctttgagcgc tattgatgta gatgttaatg 11340
atttttaggg ctgatgacca aaccaaagat gaacatggga acgnaaggaa ggttactgaa 11400
agtgtatagg cccctagttt agtcttcagt gactaatgat aatatatatt attgtgacta 11460
acaagtgttt tatagaaaca nggaaagtta gatcacaata atagatatga tcaggattat 11520
tatgtggtac ccatcccttа ttgatgaaaa tcaatggttg gttctcatag gataatcgaa 11580
aaggttaagg atcaactgta aatggagttg ttggacactt agagtagtga tttgaccttt 11640
tttctttggt agtactataa acggacatga aatgcgtagc tttacctaaa caagtctagt 11700
taagtatgat gatgcacact tgtgaatact agtgctaggt aaacccatga gatctcatgt 11760
gaagttcgaa acaaaaccta attcgaaaag tgattaaaac atgtgactta acaatgttgt 11820
agtagcattg gtcgagtttg atgggcaccc gatatgggtc actagacatg agtgtgccct 11880
gttgtgtttg agtgaagcac tagcatatca ggtgtgcaac agatatggtg cacccaggca 11940
ggacacccaa agagcttgca aaattagcct aaaacactta gtgctcacca gacatatcta 12000
gtgtactact agttattctc gttatatatg aaccctatta gttattcttg aattgcttcg 12060
atcttttaca aaggaagtag tttttccttc atctccataa actgtggttt tccaaaggca 12120
ttaataataa gatttagtat attaaattca aagttgaggt actttattat cgtgaaacca 12180
acattaatac tatagactta actaaggagt ctattggtgc ttccttctca tgtattttct 12240
tcttgaagtg ttccttcatc ttggtgctaa cgacgacatt caacaatgtg tgctcttact 12300
tgattggttt gtatatatgg tggtgttcct ttacttagtg gcaacatacc ttatcgataa 12360
ctaacccttа gtgaaagaaa tgaaaatgta catcccactg ggaaatcact catacccctа 12420
agagctaact taatggaaca tcactcatag ccctaagggc tagttggaag tactttctca 12480
tttcctgtat aagggctagt tcatgattca acttcttctc catttcttgg tgaactatct 12540
tagcacgatt cctataaaaa catatacaac taaacaaagg gtggtggtac tgaacacagt 12600
ggacccaagc actcggaaat gggaaggaca agttgcatgg aaaaaacgac aggctgggaa 12660
ctattgtgtc ttgtcaagcg tgttcgtcca gctataggac atgggtattt atagggcaac 12720
tagaggttgg tatcctaaaa tatgtccaga cccctagtta tcaactacgt tcctagataa 12780
tactgtacaa caaggtaatt atagaatagt aagtttgtta ttctaactcc accccgacag 12840
gtgggtccgt tgtcgcccgg ttgagagtgg gccctgctcg gccaggtcat tggcattgtc 12900
cgtgcagacg tgttcccaat atcgaggcaa tgaagttgtt tgacacttct tcgggagtcg 12960
gcgtgaggcc ttcgcttgct agcgcgaact tgcccacgag cgtcctcacc atgggccccg 13020
ctgacaagct t 13031

<210> 6
<211> 11232
<212> DNA
<213> Zea mays

<220>
<221> unsure
<222> (11155)...(11232)
<223> N = A, T, C, or G

<400> 6
tttttcacac cgttactgtc atctaacaga agcaggtaca aacttgtttt tcgttttcaa 60
gtcgaatttt gaggggcaaa ccatagttgc acttccatcg agggacaaaa acacaattgc 120
cccttaactt atatagttaa atatagttaa cgagcttgct actgagacta acaagtcaaa 180
actattggct tgaccttata ttagttttgt cttacacttt acaatcgttg atggctgctc 240
tagatcttat aaacttaaga atattatgac tttatcactt tatttgtaat ggatgtatgg 300
atactcattg atgcattatt tatggtataa actatagacc atgaatgtat ggtgtaatgc 360
tatagtatat tgttagactt gtgtacatat atattattta tacttaactc acaaacttaa 420
tgagtcagct cgaacttata aacgacctga gtcgacctgg ccttatggct tgttaagata 480
acaagtcaaa ccaagccgaa ctgactcgtt atccaaatct acacttacat aaacaaaaca 540
tgatttcaaa ttaagattgg tacaaaagtg ttttgtttta ttcaattaaa ccctacactg 600
tactctttat gtcaacaata gttgatgcta cgacaaagca atgaacattt tatggagtag 660
ttaattttat tgtcctaatg tcaattacta ttgttagcca aggaatggag taagccaata 720
aagagtacat atctacgagg aaatttagat atgtgcgtaa ctttttttaat cgagatacaa 780
aatgtgcaaa ataagggtcc atgtaacata catatatttc ttgtttttat ggtaaaagag 840
tgtataaact ataaaggttg ttgcttagaa gcgggattta ataacatcgg ttttatatta 900

accttaagtc cctatgcaat acctgtattt ttttctaagt acatggtaca aacacaaata 960
cacacattta agcacacata ctcacttgct atgagcacac acacgtaaac cctactccta 1020
ctagcacctt caaaagacaa aatagataaa tcttgttgac aaagtctatt gaaaaatatc 1080
aacgtccggt ctaaatcttg acaaaatatt agcacttgtg ccaagttaag aagtgagcac 1140
ttgaacgtaa gtggttagag gaacctaacc aagttagtta tgttcaattt ttcatgcaag 1200
ttagcttgct agtttttcta tacacaaaca ttatattagc ttataccatt gttgggaaat 1260
tctaacttta atgatttctt tgagaaatcc ataagagcga taaagaggag agagagagag 1320
agcaagagat ttgtacatgt ataaatacta tccattttct atttaagaat ctagacaaac 1380
tagcaaatat aaatttgaaa cataataaag atgggcacct ggcatctcct ggatattaaa 1440
agcgtaccat taaagatata cataattatt cacctcttct aggtataaat taccctacta 1500
ccacattccc ctatctctac aaactctctc tcattgactc atcaagagag tgccacctct 1560
atctctcctt ctctcttttc aaatgttcta caattatcaa ccatcataca acattcacct 1620
ttcctaccaa ccttgttgat gcttgtctca actttctctt tacctagatc actcatatat 1680
atccctattt caaaggcatt aatcatcaaa aacctataga aaaatcccat tatcaaccat 1740
gatggagtct gatcgtgaga aacaacagtc tcatggcaag aaacaaggtg accatggtag 1800
caagatgcat gattctgatg gcaataaaaa tgtgtcagat gaaaagagtc aagagtctgg 1860
tggtaaggaa cacaaatcca atataaagaa acatgaatca cgtagaaaga ggtaagacat 1920
tctccttgaa aatcttggct tcaaactcaa gttaaattta tgtacacatg tttatataga 1980
gtctagagat tttgtgctta atatatgcat gcacatgagt tcaaataatt tcataataaa 2040
aataaaaaaa tcaatatgat caggaattaa accatgaaat ttttagagac atcatctaga 2100
ttgagttcca tggtcatacc atgatggtta tgtcatttct ttccaatata aaaaattcct 2160
taacttatac tcaaaatgtt gattggatgg aacttttttct atagaattcc ttgccacatg 2220
ttgtgtaaca accatttgta ttggtttgcg tctagtccac ttttgtgtgt tgctattatg 2280
taaataatta tttttcaaat ccaaagttgt tcctccacat atctagaata tattctaatt 2340
ctacaagaat ttaaaatgaa ttgttaactt aagaatgcat tgttcaatat atttatgcat 2400
tttctcccat tatgatatat atattctcaa tatttggcac ataataactt ggaacattcc 2460
ttacatttgt tgggttgagt gctatatgtt tggattcatt aattatttac attgatattt 2520
ttgtagatgt ttgtgtttac ccaataagaa aaggccatta agaaaataaa atgttattag 2580
atagagttag tcttgacatg ttatattctt ttaataattg gattttgtgg tatttccaac 2640
acattccttc catttaaacc taactccatc tctcttatct tcctctatca tataccttat 2700
cttctttcta cactaacact aatgcttatg tcactcctaa ccttgatgca acctaccaat 2760
agtcaattac tgttacgttg ctagaaccaa agattggtcc attggtgcac aatccattag 2820
ttcctccttc ttgggactct tcaaccatcc taactcccca aatgatttca aaagttttcc 2880
ctaccatgtc atcctactcc atatccaatg tctactggtg ctagattcta tctactgtta 2940
gcaccaaact aaccacaaaa taataatccc tacaaatata ggtggaggtg atgtaaaatt 3000
aagggagggg caattgtaaa tggtagtacc atagatatca aaccttctca acttagagct 3060
atgtctacat agttctagtc ctatgaagca tcaaccattt tcttactaaa ctaaatattt 3120
ttagaggaag gggtggatcc ttactttcat ctccatgagc ttccaccccct tcctatgagc 3180
ttatccatcg actgaaagtt cctcattgct ggagcttacc cgttattatc ccatgtcatc 3240
tgacttttgt atgtactatt atctttgaag tcgtaggcat gtggtaaatt cctaccttaa 3300
gatccattaa tcctccaaca cacccttaag acccaaacca taacgcctaa atccaatttc 3360
aacatatttt aggtgacatg ggtatatgtg atattagtta cttaatatag caagctctat 3420
caatgatttt tagtcagaaa atggttgata tgtttttagt ggttgtacta taattgaaga 3480
ggcacataga gcaagttttt agaccatgaa tatatggtgt aaactataga ccatgaatgt 3540
atggtgtaat gctatagtat attaattatt agacttatgg acatatatat tatttatact 3600
taactcacaa acttaataag tcagctcgaa cttataaacc acctgagtcg aactggcctt 3660
atggctcgtt aagctaataa gtcaaaccaa gtcgagctga ttcattatcc aaatctacac 3720
ttatgtaaac aaaacatgat ttcaaattaa gattggtaca aaagtgttct gttttattca 3780
attaaacgct acactatact ccttatgtca acaatagttg atgctacgac aaagcaatga 3840
acattttatg gattagttaa ttttattatc ctaatgacaa ttactattgt cagccaagga 3900
atggagtaag ccaataaaga gtacatatct atgaggaaat ttagatatgc gtgcaacttt 3960
attttttaa tcgagataca gaatgtgcaa aataagggtc catgtaacat acatatattt 4020
cttgttttta tggtaaagga gtgtataaac tataaaggtt gttgcttaga agcgggattt 4080
taataacatc aattttatat taaccttaag cccctatcca atacatgtat tttatttcta 4140
agtacctggt acaagcataa atacacacat ttaagcacac atactcactt gttatgagca 4200
cacacgtaaa ccctactcct actagcacct tcaaaagaca aaacagatag atcttgttga 4260
caaagtctat ttatggtata aactatatac catgaatgta tggtgtaatg ctatagtata 4320
ttgttagact tgtgtacata tatattattt atacttaact cacaaactta ataagtcagc 4380
tcgaacttat aaacgacccg agtcgaactg gccttatggc tcgttaagat aacaagtcaa 4440
accaagccga gctgactcat tatccaaatc tacacttata taaacaaaac atgatttcaa 4500
attaagattg gtacaaaagt gttctatttt attcaattaa accctacact atacacctta 4560
tgtcaacatt agttgatgct acgacaaagc aatgaacatt ttatggatta gttgatgcta 4620

caacaaagta tattgttaga cttgctagat tctatctact gttagcacca aactaaccac 4680
aaaataacaa tccctataac tataggtgga ggtgatgtaa aattaaggga ggggcaattg 4740
tatatggtag taccatagat atcaaacctt ctcaacttag agctatgtct acatagttct 4800
agtcctatga agcatcaacc attttcttat actaaactaa atatttttag aggaaagggg 4860
tggatcctta ctttcatctc catgagcttc caccccttcc tatgagctta tccatcggtt 4920
gaaagtttct cattgctaga gcttactcgt tattatccca tgccatctga cttttgtata 4980
tgtactatta tctttgaagt cgtaggcatg tgtaaattcc cacctcaaga gtcaagatcc 5040
attaatcctc caacacaccc ttaagaccca aaccataaca cctaaatcca atttcaacat 5100
attttaggtg acatgggtat atgtgatatt agttacttaa tctagcaagc tctattaatg 5160
atttttagtc agaaaatggt taatatgttt ttagtggttg tactataatt gaagaggcac 5220
atagagcaag tttttagtcg ttgtattcta aacaatgatt gatgtgtata aatttaataa 5280
attcattgtt gcatcttgtg tttcatacat ttgaaatgct ttgtgcctaa tctatatgga 5340
tgaagaagta aatccttcta aacttttcct tccctgcaat ctttttaaac acactctaaa 5400
ccccaaatat ctaatcctaa cctctaaacc tgatttaaat tttctaatct agtccatttg 5460
tagtgctttt atatttagtc catttgcctt atgtgcctct tgtgtataaa tagcgtagag 5520
ttctgtataa tagtcaacaa gttttgcctt ttgttgtcgg atccattttc aatccttttg 5580
tctagttcac ctattgttgt tgtgaaaaaa atgtcacaca ttttttactt ccccctatac 5640
cacatactcc atcacggact aatgatcttc aaggtatgta tgctcagttt aaatccatgt 5700
ctccacatac tccatcttaa gttcaagtct ctactttaag gtatgtaatt ttaaaacttt 5760
gacgtattgt aattctataa ggagcaaatc tgaaaattaa ataaggaaaa actggtaaag 5820
gcatgtttgg aaatcggaac gcagacattt tgttgttcct atgtttttct ttaaataaac 5880
tcattcgtgt aaaatttctt caaaattcct ctccttcgaa cagatccttt tgcccccgga 5940
ccccttttct acgcttgccc aaacccacaa aaccctcgcc gtcgcgccgc gcgattgcct 6000
ctccggccgc cgcgagcccg cgacactagt aacggtctac accaccagaa tgactgaaga 6060
attgaattcc agcaaattca agcttttgtt ttagccaaga tttgagattc gatttgaagt 6120
gtggaagtcc ttccaatttg ccaatcctat atttgatctc tgctgtgctg cgttaaatcc 6180
ctaaacttca cagcgcggcg ccggcccagc cacgccggaa gaggtcgccg cgtgaggtca 6240
gtgtccccgt tgctgccgcc tctaacccga agcctaggcc gctgccggtg cataacaagg 6300
agaatcaggc ggaggggaaa gtagcagagg agggggcagc aactgaggag ggggagaagt 6360
accgggcgga accggaaatc ttgccgctgc cgccggccat ggcgaagctg gaagcttttg 6420
ttttagccaa gatttgagat tcgatttgaa gtgtggaagt ccttccaatt tgccaatcct 6480
atatttgatc tctgctgtgc tgcgttaaat ccctaaactt cacagcgcgg cgccggccca 6540
gccacgccgg aagaggtcgc cgcgtgaggt cagtgtcccc gttgctgccg cctctaaccc 6600
gaagcctagg ccgctgccgg tgcataacaa ggagaatcag gcggagggga aagtagcaga 6660
ggaggggca gcaactgagg agggggagaa gtaccgggcg gaaccggaaa tcttgccgct 6720
gccgccggcc atggcgaagc tgggcccggg gcaggggctc gggtgcgagg cggcggaggg 6780
gtcgctcgtg cccagccgga agcgggagta ccaagccctg cggcaagcac actgagggga 6840
agcgcccgct atatgctatc gggttcaact tcatggacgc gcgctactac gacgtcttcg 6900
ccaccgtcgg cggcaaccgc gtaagccatc gactgctctc tcctgtcgtc ctttttttgt 6960
ttctactgag gtttggggag ttcttgttga ttaatggcaa ggtaaaacta cgttgttttt 7020
ttttgtgatt ttggtggtcg gttttaggaa gcggtcgctt ttgattcaaa tttgatctaa 7080
agctgaggca ttcggttgtt tttattgggg acttgaggtc tgtaatgttc cgactattgt 7140
gatttgtttt gccgaaacat ggagtttgct agttcatttg atgaaaagct gcaacctttg 7200
acaaagaatt tgtatcactt gggaaagtat agtgaggtgt ggggaatcag atagtaccaa 7260
tattactttg actatgatta taagataatc ttttaatgtc ctttgtaacg accatgctgc 7320
ttttcgctta tcttgcctat tgatcttgca ggtgacaact taccgctgcc ttgagaatgg 7380
tagtttcgct cttctacaag cttacgttga tgaggatgta agaaagacaa tgctcaatga 7440
caatgctttt gcttgctgat ttaatattga taatattctt tctctaattc ttgtgacgcc 7500
tatttacctc agaaggatga gtcgttctat actctaagct gggctcgtga ccatgttgat 7560
ggctcaccac tgctggtggc agcaggaagc aatgggatca ttcgggtcat caattgtgct 7620
acagaaaagt tagctaaggt aatctaccct tatatttgta tgtgttccta tggtaaactt 7680
gaatgaagcc ttatttgcat aattcaatat ttcagttgtt tatttgacat atatcacttt 7740
atttatgata tctgatccag aaggtctttt ggatttgctt tagttaagga atggtgcttg 7800
ctacgcatta ataccataag caaactgtac cttttgctca cagaatattg ttaattttga 7860
ctacttcagt atgtccgttg tagtaaaaac aaatcaactt ggtgtatcta ttttttcctt 7920
gcttatacat agccaggaga ttgggcatgt ggcatgtcaa taaatactat cctataccat 7980
ttgataggac acgcactgtg tcttatttgg tagctctgtt tacgtgattc tgcagagctt 8040
tgttggccat ggcgactcaa taaatgtgat aagaactcaa ccgttgaagc cttcgctcat 8100
catttctgca agcaaggtta tgcgatagtc tgttcttagg ttcatgtacc tttttatttt 8160
tataatcttt ctgaattttg acaccatttc atatggcatt atctaatagg atgaatctgt 8220
taggctatgg aatgtccata cagggatctg tatcttgata tttgctggag ctggaggtca 8280
tcgcaatgaa gtattgagtg ttgtaagtag tgcctgctat tatgacattg tgcccttcaa 8340

aaaaaacatt attatgacat tatttttaga acattactag gttaaggtgc ctttaatatg 8400
gcgcactctt tcagctcctg atattaccat ttgttattga gcgttacatc agagataaaa 8460
taaggctacc taatgactgc tactgctttt gtactttgat tacattagtc ataaatgtac 8520
tgatgaatac attattttgt cttaaggact tccatcctag tgatattgaa cgttttgcaa 8580
gttgtggcat ggacaacact gtgaaaatct ggtcaatgaa aggttagaaa gctacttcaa 8640
agttgcttca tatttgcatg ttgcgtgtca ttgagttcac caatgttgtc gcagaatttt 8700
ggctatatgt tgacaaatca tattcatgga ctgaccttca tcaaagttcc acaaaatatg 8760
gccagtttcc agtatgtttc acaatgccta tatccaatta tcctggcaag gtcctgttgg 8820
tgtctaatcc tcatgccatc agactgacct gtttcttttt gtttcaggtc ttgattgctg 8880
cagtacactc taactatgtt gattgaacaa gatggcttgg tgacttcatc ctatcaaagg 8940
tgaaatttct gattcgttta aatggataca aatttctgta gcacggttgt cactcttttg 9000
tgggtttgac atgccactgt cttggttcat ctattgctgt accgtgcaag tgttcagttt 9060
tttcaatctt ttttctcagt gcttaatgag gggagattct atttgcagag tgttgtcaat 9120
gaaattgtgc tttgggaacc gaagacaaaa gaacagagtc ctggggaggt aattcagttt 9180
aactttccca gaattgtatt cctattataa tgccatatat ttacgcacag ttgtaaacta 9240
tttccagatc cttagatttc aaggtactgg ctgccaatat taaatatgtt ccactgaagt 9300
aatatgattt tctgttgcct catagggaag catcgatatc cttcagaagt atcctgtccc 9360
agaatgtgac atttggttta tcaaattttc atgtgatttt cacttcaatc agttggcgat 9420
aggtaatatc tctcatcagg attgtttctg gtagaagttt tatttaagat ttttttgct 9480
ctgtaaaatt tcacacacgc acacatgcac ccccacacac acacacatgc acgcacaccc 9540
ccacccacct gcacgcgcgc gtacacacac accgcacaca tatatatgac ttttttcccc 9600
acacaaatat ttgctgtgtg agatatcagc aaataaattc gtatgtttga ttatattcag 9660
agatatagga aaattgagtg ctctaatacc ccatccacta cttcaaacag gcaaccgtga 9720
aggcaaaatc tacgtgtgga aaaatacagt ccagccctcc tgtcctcatt gctcggtagt 9780
tttcactgga agagtttcag ttattcttgt ctcccacttg tatcgtcgca tgcttctgga 9840
tgccaatgct tcatcatttt caggctgtat aatcagcagt gtaaatcgcc gataagacaa 9900
actgcagtgt ccttcgatgg aaggtacctc actctaatcc atgctcaatt tggtgtactg 9960
tctattctag cacttgcttt tttcttggtt ctgcttgaga aattctcgat tgcatgtcat 10020
atgctggtgc attttctttt ttctgtttcc gtggcggatt ggtaaaatgc gacgatgcct 10080
tccttatcta gcacaatcct tggagctggt gaagacggca ccatctggcg gtgggatgaa 10140
gtggaccatc cgagctccag aaactgaaga agtgttgccg ctcaatgctg gactgatggt 10200
tacgctcggt tggggttgtg atggttgaat ccgttggcgg aaagtgccac ctggtgtttt 10260
tttctagtca aaatggttga tgttaacaga atattgaatg cttcgaatgt tgaaagttgg 10320
gatgcttgtg ctggtactct gctccgcgga cgagtgaact tagtttgttg caactttggg 10380
aaccgttgtc atctgtttgt tctgcatttc taaaaagaga gcaaatttca ggatacatgt 10440
tctttttttt cagtacagga aaactaaggt tgaggtattg ctttgcaatt tactctctct 10500
ctctctctct cttaaaaaaa ctggatcttg cttcaacgat gcattccttg ggtcatcggt 10560
tttacttttg aaatcttgat agctgggcct aaagttacca agcccactag tatcagaagt 10620
aataatatga tggctcctcc cctgccttac tgtcacgtgt aaactttcga aactagcagg 10680
actgtagcat ttagcgagct ggttgtttgg gttagagctc agcgtcgcaa cttatggtac 10740
cgaggtcagt gtcaagatct atggcaccat ggttcaatca cagttttagt cccaccaaaa 10800
atataaaggt gaagtttcga caaaaaatgg ctagaataaa aaaaaacagg tccacatact 10860
gaggagaaca catgacagat tcaccaagga ttttgaattg aaagaggcta atgattgaca 10920
ggatttgatc ttcaattcca cctcccgttg tcctgcttct actctaaagt tcaagcgtgg 10980
ctcagtttgg ctatctgtta taatttcaag aaatcctgat ttctgttagc agtttactag 11040
gctattagga ggagctggga caaaagaaaa acgagaattg acgaggacaa attcgcaatt 11100
agttgggaaa ttgggggcac aattttcaat gcccacaaaa ttcactcccc ctacntntgc 11160
ggnggaatgg ggtcannnct cantgtcccc tgttnccggg acaagtntaa ctaacacatt 11220
tccnnattnn tn 11232

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 September 2003 (25.09.2003)

PCT

(10) International Publication Number
WO 03/078580 A3

(51) International Patent Classification[7]: C12Q 1/00, 1/68, C12P 19/34, C12N 5/00, 15/64, 15/82

(21) International Application Number: PCT/US03/07552

(22) International Filing Date: 13 March 2003 (13.03.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/363,861 13 March 2002 (13.03.2002) US

(71) Applicant *(for all designated States except US)*: PIONEER HI-BRED INTERNATIONAL, INC. [US/US]; 800 Capital Square, 400 Locust Street, Des Moines, IA 50309 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: DANILEVSKAYA, Olga [US/US]; 6004 Dogwood Circle, Johnston, IA 50131 (US). HERMON, Pedro [US/US]; 9814 Newport Vista Drive, Johnston, IA 50131 (US).

(74) Agents: VARLEY, Karen, K. et al.; Darwin Building, 7100 N.W. 62nd Avenue, Johnston, IA 50131-1000 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

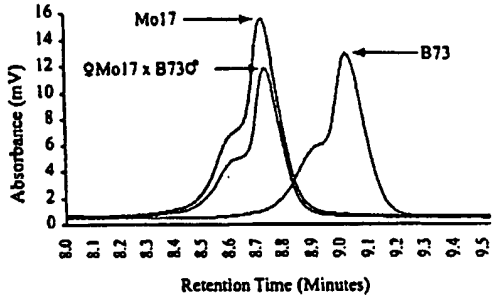
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR).

Published:
— *with international search report*

(88) Date of publication of the international search report:
27 November 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IMPRINTING IN PLANTS TO CONTROL GENE EXPRESSION



(57) Abstract: Compositions and methods for identifying imprinting and genes regulated by imprinting are provided. The methods involve an analysis of the nucleotide sequence and the identification of CpG islands. At least two islands are involved in imprinting. Thus, genes can be identified that are differentially expressed based on parental inheritance. In this manner, the methods are useful for determining the propensity of a gene to be influenced by imprinting. Such analysis involves determining the pattern of imprinting for cells of interest. Fig. 2A shows the pattern of paternal and maternal ZmFie1 allele expression in developing kernels. It is further recognized that DNA constructs can be constructed which show differential expression depending upon the parent-of-origin. To silence a paternally inherited allele, at least two CpG islands are utilized in the construct.

WO 03/078580 A3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/07552

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/00, 1/68; C12P 19/34; C12N 5/00, 15/64, 15/82
US CL : 435/4, 6, 91.2, 91.41, 468, 375

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/4, 6, 91.2, 91.41, 468, 375

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,786,146 A (HERMAN et al.) 28 July 1998, col. 2, line 57 continues to line 18 of col. 3.	5
A	LUO, M., et al. Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds. PNAS. September 2000, Vol. 97. No. 19, pages 10637-10642.	1-7
A	YADEGARI, R. et al. Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms. The Plant cell. December 2000, Vol. 12, pages 2367-2381.	1-7

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 May 2003 (14.05.2003)	05 SEP 2003
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703)305-3230	Authorized officer Quang Nguyen, Ph.D. Telephone No. (703) 308-2801

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

PCT/US03/07552

Continuation of B. FIELDS SEARCHED Item 3:
APS, DIALOG, MEDLINE, EMBASE, BIOSIS
search terms: plant, zea mays, arabidopsis, imprinting, imprinted, CpG islands, endosperm, cytosine methylation, PCR, restriction enzymes.